

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 . Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 15:54:39 ; Search time 65 Seconds  
(without alignments)  
5104.524 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1. gccactctctctctggagga.....acaaggagccccccagaag 1245

oring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL-frame+np.model -DEV-xml  
-DB-A\_Geneseq\_101002 -OFMT-fastan -SUFFIX-n2p.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=pblosum62 -TRANS-human40.cdi  
-LIST=60 -DOCALLIGN=200 -THR SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE-LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-US09763153 @CN1\_1\_81 -runat\_11032003\_084246\_1971 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_101002:\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	2298	98.3	415	21	AAV56803
2	2298	98.3	419	22	AAE08625
3	2298	98.3	419	22	AAE08625
4	2298	98.3	419	22	AAE08625
5	2298	98.3	419	22	AAE08625
6	2298	98.3	419	22	AAE08625
7	2298	98.3	460	9	AAE08625
8	2298	98.3	460	18	AAE08625
9	2298	98.3	461	7	AAE08625
10	2298	98.3	461	8	AAE08625
11	2298	98.3	461	10	AAE08625
12	2298	98.3	461	12	AAE08625
13	2298	98.3	461	12	AAE08625
14	2298	98.3	461	12	AAE08625
15	2298	98.3	461	14	AAE08625
16	2298	98.3	461	17	AAE08625
17	2298	98.3	461	20	AAE08625
18	2298	98.3	461	22	AAE08625
19	2298	98.3	461	22	AAE08625
20	2298	98.3	461	22	AAE08625
21	2295	98.2	419	23	AAU99001
22	2295	98.2	419	23	AAU99001
23	2295	98.2	419	23	AAU99001
24	2295	98.2	461	9	AAE08625
25	2294	98.2	419	10	AAE08625
26	2293	98.1	419	23	AAU99001
27	2293	98.1	419	23	AAU99001
28	2293	98.1	419	23	AAU99001
29	2292	98.1	419	23	AAU99001
30	2292	98.1	419	23	AAU99001
31	2292	98.1	419	23	AAU99001
32	2292	98.1	419	23	AAU99001
33	2292	98.1	461	12	AAE08625
34	2292	98.1	461	12	AAE08625
35	2292	98.1	461	12	AAE08625
36	2292	98.1	461	12	AAE08625
37	2291	98.0	419	14	AAE08625
38	2291	98.0	419	14	AAE08625
39	2291	98.0	419	23	AAU99001
40	2291	98.0	419	23	AAU99001
41	2291	98.0	419	23	AAU99001
42	2291	98.0	419	23	AAU99001
43	2290	98.0	419	23	AAU99001
44	2290	98.0	419	23	AAU99001
45	2290	98.0	419	23	AAU99001
46	2290	98.0	419	23	AAU99001
47	2290	98.0	419	23	AAU99001
48	2290	98.0	419	23	AAU99001
49	2290	98.0	419	23	AAU99001
50	2289	97.9	419	23	AAU99001
51	2289	97.9	419	23	AAU99001
52	2289	97.9	419	23	AAU99001
53	2289	97.9	419	23	AAU99001
54	2289	97.9	419	23	AAU99001
55	2289	97.9	419	23	AAU99001
56	2289	97.9	419	23	AAU99001
57	2289	97.9	419	23	AAU99001
58	2289	97.9	419	23	AAU99001
59	2289	97.9	419	23	AAU99001
60	2289	97.9	419	23	AAU99001

## ALIGNMENTS

RESULT 1  
AAV56803  
ID AAV56803 standard; Protein; 415 AA.  
XX

AC AAY56803;  
 DT 27-MAR-2000 (first entry)  
 XX Truncated human protein C polypeptide.  
 DE  
 XX  
 XX protein C; truncated; thrombotic disorder; vascular disorder; stroke;  
 KW hypercoagulable state; myocardial infarction; unstable angina; sepsis;  
 KW adult respiratory distress syndrome; sickle cell anemia; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W09563070-A1.  
 PN  
 XX 09-DEC-1999.  
 PD  
 XX 01-JUN-1999; 99W0-US11969.  
 PF  
 XX 01-JUN-1998; 98US-0087585.  
 PR  
 XX (ELIL) LILLY & CO ELI.  
 XX  
 XX Huang L, Riggan RM;  
 DR WPI; 2000-086975/07.  
 DR N-PSDB; AA246750.  
 XX  
 PT Novel polypeptide useful for treating thrombotic and vascular diseases  
 PT and hypercoagulation, e.g. stroke  
 XX  
 PS Claim 2; Page 22-23; 23pp; English.  
 XX  
 XX This represents a human protein C polypeptide having a light chain and  
 CC a truncated heavy chain. The protein can be produced by standard  
 CC recombinant methodologies. The truncated protein C is used to treat a  
 CC wide range of thrombotic or vascular disorders or hypercoagulable states,  
 CC e.g. stroke; myocardial infarction; unstable angina; sepsis; adult  
 CC respiratory distress syndrome; sickle cell anemia etc. The truncated  
 CC protein C retains the activity of full-length protein C but does not  
 CC undergo C-terminal cleavage, of the heavy chain, during activation.  
 XX  
 XX Sequence 415 AA;  
 SQ

Alignment Scores:  
 Pred. No.: 1 34e-159 Length: 415  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 21 Gaps: 0

us-09-763-153-1 (1-1245) x AAY56803 (1-415)

1 GCCAACTCTTCCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGTCATAGAGGAG 60  
 |||||  
 1 AlaAsnSerPheLeuGluGluLeuAArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 |||||  
 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGACACACTGGCCCTTC 120  
 |||||  
 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40  
 |||||  
 121 TGGTCCAAAGCAGCTGCGAGCGTGACCATGCTGTGCTTGGCTTGGAGCACCCTGGCGC 180  
 |||||  
 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
 |||||  
 181 AGCTGTGCTGCGGCGGAGCGAGCTGGATCGACGCGCATCGAGCTTCAGCTCGACTGC 240  
 |||||  
 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 |||||  
 241 CGCAGCGCTGGAGGCGCGCTTCGCGAGCGGAGGTGAGCTTCCTCAATTGCTCGCTG 300  
 |||||  
 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 |||||

QY 301 GACAACGGCGGCTGCACGCATTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
 |||||  
 Db 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 120  
 |||||  
 QY 361 GCGCTGGCTACAAGTGGGGGAGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 |||||  
 Db 121 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
 |||||  
 QY 421 TGTGGGAGGCGCTGGAGCGGATGGAGAGAGCGAGTCCACCTGAAACGAGACACAGAA 480  
 |||||  
 Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 |||||  
 QY 481 GACCAAGAAGACCAAGTAGATCCGCGCTCATTCATGGGAGATGACACAGCGGGGAGAC 540  
 |||||  
 Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 |||||  
 QY 541 AGCCCTGGCAGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 |||||  
 Db 181 SerProTrpGlnValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
 |||||  
 QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 |||||  
 Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 220  
 |||||  
 QY 661 GTCAGGCTGGAGATGATGACCTGCGGCGCTGGGAGAGTGGGAGTGGGAGTGGGAGT 720  
 |||||  
 Db 221 ValArgLeuGlyGlyLysAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 |||||  
 QY 721 RAGGAGGTCTTCTGCTCCACCCCACTTACAGAGAGACACCCAGCAATGATCGCACTG 780  
 |||||  
 Db 241 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 260  
 |||||  
 QY 781 CTGCACTGGCGCCCGCCCGCCCTCTCGCAGACCATAGTGCCTGCTGCTGCTGCTGCTG 840  
 |||||  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 |||||  
 QY 841 AGCGGCTTGCAGAGCGGCGGCTCAATCAGCGCGGCGGAGGAGCTGCTGCTGCTGCTG 900  
 |||||  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 |||||  
 QY 901 GGCTACACAGCAGCGGAG 960  
 |||||  
 Db 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 |||||  
 QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 |||||  
 Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 |||||  
 QY 1021 GAGAACAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
 |||||  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 |||||  
 QY 1081 GGGGGGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 |||||  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 |||||  
 QY 1141 GGTGAGGCTGTGGGCTGCTTCCAACTACAGCTGCTTACCAAGTACCAAGTACCGCT 1200  
 |||||  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
 |||||  
 QY 1201 GACTGGATTCATGGGCGCATCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1245  
 |||||  
 Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

RESULT 2  
 AAE08625  
 ID AAE08625 standard; Protein; 419 AA.  
 XX  
 AC AAE08625;  
 XX  
 DT 01-NOV-2001 (first entry)  
 XX  
 DE Human mature wild type protein C.

XX Human; protein C derivative; anticoagulation activity; thrombosis;  
 KW serpin inactivation; acute coronary syndrome; myocardial infarction;  
 KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;  
 KW disseminated intravascular coagulation; DIC; burn; transplantation;  
 KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;  
 KW haemolytic uraemic syndrome; acute arterial thrombotic occlusion;  
 KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.  
 XX Homo sapiens.  
 XX WO200159084-A1.  
 XX PD 16-AUG-2001.  
 XX 02-FEB-2001; 2001WO-US01221.  
 XX 11-FEB-2000; 2000US-0181948.  
 XX 14-MAR-2000; 2000US-0189199.  
 XX (ELIL) LILLY & CO ELI.  
 XX Gerlitz BE, Grinnell BW, Jones BE;  
 XX WPI; 2001-514662/56.  
 XX DR N-PSDB; RAD15223.  
 XX Protein C derivative for treating acute coronary syndromes, vascular  
 PT occlusive disorders, thrombotic disorders and sepsis, comprises  
 PT substitutions at specified amino acid positions -  
 XX Claim 1; Page 43-44; 59pp; English.

The invention relates to human protein C derivatives and nucleic acid  
 molecules encoding such derivatives. These derivatives have increased  
 anticoagulation activity, resistance to serpin inactivation and  
 increased sensitivity to thrombin activation compared to wild type  
 protein C, and retains the biological activity of the wild type human  
 protein C. Protein C derivatives are useful in the manufacture of a  
 medicament for the treatment of acute coronary syndromes e.g. myocardial  
 infarction and unstable angina; and disease states predisposing to  
 thrombosis; vascular occlusive disorders and hypercoagulable states e.g.  
 disseminated intravascular coagulation (DIC), burns, transplantations,  
 thalassaemia, sickle cell disease, viral haemorrhagic fever and  
 haemolytic uraemic syndrome; sepsis in combination with bacterial  
 permeability increasing protein; sepsis in combination with bacterial  
 with an anti-platelet agent; thrombotic disorders in combination  
 thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral  
 or peripheral arteries or in vascular grafts in combination with a  
 thrombolytic agent. Nucleic acid molecules of the invention are useful  
 for treating humans with genetically predisposed prothrombotic disorders  
 by gene therapy. The present sequence is human mature wild type  
 protein C.

SQ Sequence 419 AA;

Alignment Scores:  
 Pred. No.: 1,34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 Dbs: 22 Gaps: 0

US-09-763-153-1 (1-1245) x AAE08625 (1-419)

QY	121	TGTTCAAGCAGCTGACGCTGACGCTGCTTGGTCTTGGCCCTTGAGACACCCGTCGCC	180
Db	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuLeuGluHisProCysala	60
QY	181	AGCCTGTCTGCGGGCAGCGACGCTGATCGACGCGCATCGGCGATCGGATGTC	240
Db	61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
QY	241	CGCAGCGCTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCCAATGCTCGCTG	300
Db	81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
QY	301	GACAAAGCGGCTGACGCTGACGCTGCTGAGGAGTGGGCTGGCGCGCTGTAGCTGT	360
Db	101	AspAsnGlyGlyCysThrHisTyrcysLeuGluValGlyTrpArgCysSerCys	120
QY	361	GGCCTGTCTACAAGTGGGCGACCTCTCGAGTGTACCCCGCAGTGAAGTTCCT	420
Db	121	AlaProGlyTyrcysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro	140
QY	421	TCTGGGCGCTGGAAGCGGATGGAAGAAGCGAGTGCCTGAAACGAGACACAGAA	480
Db	141	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgThrGlu	160
QY	481	GACCAAGAAGACCAAGTAGATCCGCGCTCATTCATGGGAGATGACCGCGGGAGAC	540
Db	161	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	180
QY	541	AGCCCTTGGCAGGTGGTCTGCTGACTCAAGAAGAAGCTGGCTGGCGGCGAGTGC	600
Db	181	SerProTrpGlnValValLeuLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTGGTGGTGTGACGCGGCGCTGATGATGATGATGATGATGATGATG	660
Db	201	IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLeuLeu	220
QY	661	GTCAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
Db	221	ValArgLeuGlyGlyLysTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAsp	240
QY	721	AAGAGGTCTCTGCTCCACCCCACTACACAGACACACACACACACACACACACAC	780
Db	241	LysLysValPheValHisProAsnTrpSerLysSerTrpThrAspAsnAspLeuAlaLeu	260
QY	781	CTGCACCTGGCCAGCGCCCGCCCTCTCGCAGACCATAGTGCCTGCTGCTGCTGCTG	840
Db	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	280
QY	841	AGCGGCTTGGCAGCGGCGGCTCAATCAGCGCGCGCGAGAGACCTCGTGCAGCGGTGG	900
Db	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACACAGCAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	960
Db	301	GlyTyrcysSerArgGlyLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAAGATTCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1020
Db	321	IleLysIleProValProHisAsnGlnCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGAACATGCTGTGGCGGCTCTCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGG	1080
Db	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GG	1140
Db	361	GlyGlyProMetValAlaSerPheHisGlyThrPheLeuValGlyLeuValSerTrp	380
QY	1141	GGTGAGGCTGTGGGCTCTTCAACTACGCGGTTCACCAAGTCAGCGCTACCTTC	1200
Db	381	GlyGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu	400

XX

1	QY	GC	AACTCCTTCTGGAGGAGCTCGTCACACAGCAGCCTGGAGCGGGAGTGTCATAGAGAG	60
1	Db	AA	ASNSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	20
61	QY	ATCTGTGACTTCTGGAGGCGCAAGGAAATTTTCCAAAATGTGGATGACACACTGGCCCTTC	120	
21	Db	IL	CysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	40
121	QY	TGCTCCAGACAGCTGCAGCGGTGACAGTGGTCTTGGCCCTTGGAGCACCCGTGCCGC	180	
41	Db	TR	SerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	60
181	QY	AGCCTGTGCTCGGGGCACGGCAGTGCATCGACGCATCGGCAGCTTCACGTCGCACTGC	240	



Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 QY 241 CGCAGCGGCTGGAGGCGGCTTCGCCAGCGGAGGTGAGCTTCCTCAATTCCTGCTG 300  
 Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 QY 301 GACACGGCGGCTGACACCATTTACTGCTAGAGGAGTGGCTGGCGGCTGTAGCTGT 360  
 Db 101 AspAsnGlyGlyCysThrHisTrpCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
 QY 361 GCGCTGCTACAGCTGGGAGCAGCTCCTGCAAGTGTACCCCGCAGTGAAGTTCCT 420  
 Db 121 AlaProGlyTrpLysLeuGlyAspSerLeuLeuGlnCysHisProAlaValLysPhePro 140  
 QY 421 TGTGGAGCGCTGGAAGCGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATTGATGGGAAGATGACCGAGGGGAGAC 540  
 Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 QY 541 AGCCCTGCGCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
 Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
 QY 601 ATCCACCCCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 660  
 Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGlySerLysLysLeu 220  
 QY 661 GTACAGGCTTGGACATGATGACCTGGCGGCTGGGAGAGTGGAGCTGGACCTGGACAT 720  
 Db 221 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 QY 721 AAGGAGGTCTCTGTCACCCCACTACAGCAAGAGCACCACCGCAATGACATCGCATG 780  
 Db 241 LysGluValPheValHisProAsnTrpSerLysSerTrpThrAspAsnAspIleAlaLeu 260  
 QY 781 CTCACCTGGCCAGCGCCGCTTCGCGAGCAGCATAGTGGCCATCTGCTCCCGGAC 840  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 QY 841 AGCGGCTTGCAGAGCGGAGTCAATCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 900  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnThrLeuValThrGlyTrp 300  
 QY 901 GCTACCCAGCAGCGGAG 960  
 Db 301 GlyTrpHisSerArgGlyGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 QY 961 ATCAAGATTCCGTTGGTCCCGCACAATGAGTGCAGCGAGGTCATGAGCAACATGGTGT 1020  
 Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 QY 1021 GAGACATCTGTGTGGGGCATCTCCGGGACCGGAGGATGCTCCGAGGCGGACAGT 1080  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 QY 1081 GGGGGGCGCATGTGCTCTCTCCAGCGGACCTGTTCTGCTGGGCTGTGTGAGCTGG 1140  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 QY 1141 GGTGAGGCTGTGGGCTCTTCAACTAGCGGCTTTACACCAAGTACAGCGCTACCTC 1200  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpLysValSerArgTrpLeu 400  
 QY 1201 GACTGGATTCATGGCACATCAGACACAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415  
 RESULT 4  
 AAB36894  
 ID AAB36894 standard; Protein; 419 AA.

XX  
 AC AAB36894;  
 XX  
 DT 26-FEB-2001 (first entry)  
 XX  
 DE Human protein C derivative 1.  
 XX  
 KW Protein C; human; vascular occlusive; burn; transplantation;  
 KW deep vein thrombosis; sickle cell; thalassemia;  
 KW thrombotic disorders; myocardial infarction; angina; stroke.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200066754-A1.  
 XX  
 PD 09-NOV-2000.  
 XX  
 PF 13-APR-2000; 2000WO-US08722.  
 XX  
 PR 30-APR-1999; 99US-0131801.  
 XX  
 PA (ELIL ) LILLY & CO ELI.  
 XX  
 PI Gerlitz BE, Jones BE;  
 XX  
 DR WPI; 2001-007227/01.  
 XX  
 DR N-PSDB; AAC83311.  
 XX  
 PT Protein C derivatives, useful for treating vascular occlusive disorder,  
 PT hypercoagulable state, thrombotic disorder and disease states  
 PT predisposing thrombosis, comprises specific amino acid substitutions -  
 XX  
 PS Claim 1; Page 42-44; 57pp; English.  
 XX  
 CC The present invention relates to a human protein C derivative. The  
 CC protein is useful for treating vascular occlusive disorders.  
 CC hypercoagulable states such as sepsis, disseminated intravascular  
 CC coagulation, purpura fulminans, major trauma, major surgery, burns,  
 CC adult respiratory distress syndrome, transplantation, deep vein  
 CC thrombosis, heparin-induced thrombocytopenia, sickle cell disease,  
 CC thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic  
 CC purpura, and hemolytic uremic syndrome, and also useful for treating  
 CC thrombotic disorders and acute coronary syndromes such as myocardial  
 CC infarction, unstable angina, and stroke. Protein C derivatives with  
 CC amino acid substitutions result in increased resistance to  
 CC inactivation by serpins when compared to wild-type activated human  
 CC protein C. They also have longer half-lives in human blood and hence  
 CC require either less frequent administration and/or smaller dosage  
 CC than wild type human protein C for treating disorders.  
 XX  
 SQ Sequence 419 AA;

Alignment Scores:  
 Pred. No.: 1.34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 22 Gaps: 0

US-09-763-153-1 (1-1245) x AAB36894 (1-419)

QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTCACAGAGCTGGAGGGAGTGCATAGAGAG 60  
 Db 1 AlaAsnSerPheLeuGluGluArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGCACACACTGGCCCTC 120  
 Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspSerThrLeuAlaPhe 40  
 QY 121 TGTGCCAAGCAGCTCCAGCGTGACACAGTCTGTGGCTTGGCCCTTGGAGCAGCCCTGCCGCC 180  
 Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60

Tue Mar 18 16:19:30 2003

## RESULT 5

AAU99002 standard; Protein; 419 AA.  
 AAU99002; (first entry)  
 Human Protein C zymogen protein.  
 Human; Protein C; N-glycosylation; APC; activated protein C; zymogen; serum half-life; chromosome 2q13-q14; stroke; myocardial infarction; after venous thrombosis; disseminated intravascular coagulation; DIC; sepsis; septic shock; embolism; pulmonary embolism; burn; pregnancy; bone marrow transplantation; major surgery; trauma; ARDS; coagulant; adult respiratory distress syndrome; alpha-1 antitrypsin.  
 Homo sapiens.  
 Key Location/Qualifiers  
 Protein 1..155 /label= Light\_chain  
 Peptide 156..157 /label= Lys\_Arg\_dipeptide  
 Peptide 158..169 /label= Activation\_peptide  
 Protein 158..419 /label= Heavy\_chain  
 WO200232461-A2.  
 25-APR-2002.  
 15-OCT-2001; 2001WO-DK00679.  
 18-OCT-2000; 2000DK-0001560.  
 18-OCT-2000; 2000US-242268P.  
 21-JUN-2001; 2001DK-0000970.  
 21-JUN-2001; 2001US-300154P.  
 (MAXY-) MAXYGEN APS.  
 (MAXY-) MAXYGEN HOLDINGS LTD.  
 Andersen KV, Pedersen AH, Freskgaard PO;  
 WPI; 2002-489875/52.  
 N-PSDB; ABK86039.  
 Novel conjugate useful for treating or preventing septic shock, stroke and myocardial infarction, comprises non-polypeptide group covalently attached to protein C polypeptide comprising an attachment group -  
 Claim 2; Page 79-81; 92pp; English.  
 The invention relates to a conjugate (I) comprising at least one non-polypeptide moiety (II) (e.g. an N-glycosyl group) covalently attached to a protein C polypeptide comprising an amino acid sequence which differs from that of a parent protein C polypeptide (III) in at least one introduced and/or at least one removed amino acid residue comprising an attachment group for the non-polypeptide group (e.g. an N-glycosylation site). Also included are (1) a variant (IV) of (III) comprising a substitution in a position (P) where (P) is an amino acid with at least 25% of its side group exposed to the surface, with the proviso that the substitution is not Thr245Ser/Ala/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln, Tyr302Ser/Ala/Thr/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln or Phe316Ser/Ala/Thr/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln; (2) a nucleotide sequence (V) encoding (IV); (3) an expression vector (VI) comprising (V); (4) a host cell (VII) comprising (V) or (VI); (5) increasing (M2) the functional in vivo half-life of the serum half-life of a parent protein C polypeptide. The conjugates, variants and protein C proteins are useful as medicaments, and in the manufacture of medicaments for the treatment (and diagnosis/prevention) of stroke, myocardial infarction, after venous thrombosis, disseminated intravascular coagulation (DIC), sepsis, septic

181 AGCTGTGCTCGGCGACGCGCATCGATCGAGCGCATCGGCGCTCGAGCTCGACTGC 240  
 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 241 CGCAGCGCTGGAGGGCGCTTCTGCGAGCGCGAGCTGAGCTTCTCAATCTCGCTG 300  
 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 301 GACACGGCGCTGCACCATCTACTAGAGAGGTGGCTGGCGGCTGAGTGT 360  
 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 120  
 361 GCGCTGCTCAAGCTGGGGACGACCTCTCGAGTGTACCCCGGAGTGAAGTCCCT 420  
 121 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
 421 TGTGGAGGCGCTGGAGCGGATGAGAGAGCGCAGTCACCTGAGACGAGACAGAA 480  
 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 481 GACACAGACCAAGTACGATCGGGCTCATTTGATGGAAGATCACCGCGGGGAGAC 540  
 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 541 AGCCCTGGCAGGTGCTCTGCTGGACTCAAGAAGAGTGGCTGGCGGCGAGTGC 600  
 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200  
 601 ATCCACCTCTGCTGGTGTGTCAGCGCGCCCTGATGATGATGATGATGATGATGAT 660  
 201 IleHisProSerTrpValLeuThrAlaIleHisCysMetAspGluSerLysLysLeuLeu 220  
 661 GTCAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 221 ValArgLeuGlyGlyLeuArgLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 721 AAGAGGTCTGCTGCGCCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260  
 781 CTGACCTGCGCCGCGCGCGCCCTCTCGCAGACCATAGTGCCTGCTGCTGCTGCTGCT 840  
 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 841 AGCGGCTTGGAGCGGCGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 900  
 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 901 GGTACACAGCGCGCGAG 960  
 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 961 ATCAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 1021 GAGAACATGCTGTGCGGGCATCTCGGGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 1081 GGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 1141 GGTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
 1201 GACTGGATCATGGGACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

CC shock, emboli e.g. pulmonary emboli, transplantation such as bone marrow  
 CC transplantation, burns, pregnancy, major surgery/trauma or adult  
 CC respiratory distress syndrome (ARDS). The variant protein C has an  
 CC increased resistance to activation by e.g. human plasma and alpha-1  
 CC antitrypsin. The conjugates have an increased in vivo half-life,  
 CC increased serum half-life, increased resistance to inhibitors, reduced  
 CC renal clearance, reduced immunogenicity and/or increased bioavailability.  
 CC The conjugate offers a number of advantages over the currently available  
 CC APC products, including longer duration between injections,  
 CC administration of less protein, and fewer side effects. Moreover, a  
 CC reduced anticoagulant activity is beneficial to reduce the risk of  
 CC bleeding while maintaining the antiinflammatory activity of APC  
 CC (activated protein C) conjugates. This must be especially important when  
 CC the conjugate has an extended plasma life. The gene for protein C is  
 CC located on chromosome 2q13-q14. The present sequence represents zymogen  
 CC protein C upon which the variants of the invention were based.

Sequence 419 AA;

# Alignment Scores:

Pred. No.: 1.34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 23 Gaps: 0

US-09-763-153-1 (1-1245) x AAU99002 (1-419)

QY 1 GCCAATCTCTCTCGTGGAGGCTCGTCCACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 1 AlaAnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 QY 61 ATCTGTGACTTGGAGGAGCCCAAGAAATTTTCCAAATATGGATGACACACTGGCCTTC 120  
 Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnValAspThrLeuAlaPhe 40  
 QY 121 TGGTCCAAAGCAGTGGAGGTCGACAGTGTGCTGTGCTTGGCCCTGGAGCCCGTGGCC 180  
 Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
 QY 181 AGCTGTGCTCGGGGACGGCAGCTGCATCGAGCGATCGGACGCTTCAGCTCGGACTGC 240  
 Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 QY 241 CGCAGCGGCTGGAGGCGCGCTCTGCCACCGCGAGGTGAGCTTCCTCAATTGCTGCTG 300  
 Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 QY 301 GACAAACGGGCTGCACGCATTACTGCTAGAGAGGTGGCTGGCGGCGCTGAGCTGT 360  
 Db 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 120  
 QY 361 GCGCTGGCTACAGCTGGGGGACGACCTCTGCTAGTGTACCCCGCAGTGAAGTTCCT 420  
 Db 121 AlaProGlyTrpLysLeuGlyAspSerLeuLeuGlnCysHisProAlaValLysPhePro 140  
 QY 421 TGTGGGAGCCCTGGAGGCGGATGGAGAAGACGCGAGTCACTCCTCAACAGACACAGAA 480  
 Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 QY 481 GACCAAGAAGACCAAGTAGATCCGGGCTCAFTGATGGAAGATGACCGAGCGGGAGAC 540  
 Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 QY 541 AGCCCTGGCAGTGTCTCTGCTGGACTCAAGAGAGAGCTGGCTGGGGGCGAGTGCCTC 600  
 Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200  
 QY 601 ATCCACCCCTCTGGTGTGACAGCGGCCCACTGCATGGATGAGTCCAAAGAGCTCCTT 660  
 Db 201 IleHisProSerTrpValLeuThAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220

QY 661 GTCAGGCTTGGAGAGTATGACCTCGCGCTGGGAGAGTGGAGCTGGACCTGGACATC 720  
 Db 221 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 QY 721 AAGGAGCTTCTGCTCCACCCCACTACAGCAAGAGCACCACCAATGACATCGCACATG 780  
 Db 241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260  
 QY 781 CTGCACCTGGCCCGAGCCCGCCCTCTCGCAGACCATAGTCCCATCTGCTCCCACTTC 840  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 QY 841 AGCGGCTTTCAGAGCGGAGCTCAATCAGGCGGCGGAGAGACCTCTGTCAGCGGCTGG 900  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 Db 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 QY 961 ATCAAGATTCCTGCTCCCGCAGCAATGAGTGCAGCGAGGTGCATGAGCAACATGTGTCT 1020  
 Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 QY 1021 GAGACATGCTGTGCGGGCATCTCTCGGGACCGGACCGGAGGAGTGCCTGGAGGCGCAGT 1080  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 QY 1081 GGGGGGCCCATGCTGCTCCCTCTCCACGGCACCTGCTTCTGCTGGCTGGCTGGCTGG 1140  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 QY 1141 GGTGAGGCTGTGGGCTCTCTTCACTCAACTACGCGCTTTACACCAAGTACGCGCTACCTC 1200  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerA-gTyrLeu 400  
 QY 1201 GACTGATTCATGGGACATCAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 6

AAP81104  
 ID AAP81104 standard; protein; 460 AA.

AC AAP81104;

XX 16-SEP-1990 (first entry)

DE Sequence of human protein C.

XX Human protein C; plasmid ppc 1.

OS Homo sapiens.

XX JP63263083-A.

XX 30-NOV-1988.

PD 21-APR-1987; 87JP-0096341.

PF 21-APR-1987; 87JP-0096341.

XX (FARH ) HOECHST JAPAN KK.

XX WPI; 1988-350711/49.

DR N-PSDB; AAN81408.

XX Human protein C gene - prepd. from new DNA of specified base sequence.

PS Disclosure; ; 16pp; japanese.

XX The human protein C is expressed in large amts. using plasmid ppc 1 in  
 CC E.coli K12/Om 225 (FERM P-9297).

Tue Mar 18 16:19:30 2003

XX SQ Sequence 460 AA;

Alignment Scores: 1.37e-159 Length: 460  
 Pred. No.: 2298.00 Matches: 415  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 98.33% Gaps: 0  
 DB:

US-09-763-153-1 (1-1245) x AAPB1104 (1-460)

QY 1 GCCAACTCTCTCCCTGGAGGAGCTCCGTCACAGCAGCTGGAGGGAGTGCATAGAGGAG 60  
 DB 42 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 61  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGACACACTGGCCTTC 120  
 DB 62 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 81  
 QY 121 TGGTCAAGCAGCTCGAGGCTGACAGGTGCTTGTCTTGGCTTGGAGCACCCTGCGCC 180  
 DB 82 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 101  
 QY 181 AGCCTGTGCTGGGGCAGCGGACGTCATCGACGCGATCGGCGACTTCAGTGGCGATGC 240  
 DB 102 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 121  
 QY 241 CGCAGCGCTGGAGGGCGCTCTGCCAGCGGAGGAGCTTCCCTCAATTGCTCGCTG 300  
 DB 122 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 141  
 QY 301 GACACGGCGCTGCACGATTAATCTGCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360  
 DB 142 AsnAsnGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 161  
 QY 361 CGCCTGGCTTACAGCTGGGGAGGAGCTTCCAGTGTCCCGCAGTGAAGTTCCT 420  
 DB 162 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 181  
 QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGGAGTCCAGTCACTGAAGAGACACAA 480  
 DB 182 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 201  
 QY 481 GACCAAGAAGACCAAGTAGATCCGGGCTCATGATGGAGAGATGACACGCGGGAGAC 540  
 DB 202 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 221  
 QY 541 AGCCCTGGCAGGCTGCTGCTGGAGTCAAGAGAGAGCTGGCGCTCGGGGCGAGTGTCT 600  
 DB 222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241  
 QY 601 ATCCACCCCTCTGGGTGCTGACAGCGGCCCTCATGATGATGATGATGATGATGATGAT 660  
 DB 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 261  
 QY 661 GTGAGCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 DB 262 ValArgLeuGlyGlyArgLysLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 281  
 QY 721 AAGGAGGTCTTGTCTCCACCCCACTACAGCAAGAGCACCACCGACATGACATGACACTG 780  
 DB 282 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 301  
 QY 781 GTGACCTGGCCAGCGCCACCTCTCGCAGACCATAGTGCCTATGCTCCCGGAC 840  
 DB 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321  
 QY 841 AGCGGCTTGCAGAGCGGAGTCAATCAGCGCGCGCGAGGAGACCTCGTACCGGCTGG 900  
 DB 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341

QY 901 GGCTACACAGCAGCGAG 960  
 DB 342 GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361  
 QY 961 ATCAAGATTCCCGTGGTCCCGCACATGAGTGCAGCGAGGTGCATGAGCAACATGGTGTCT 1020  
 DB 362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381  
 QY 1021 GAGAACATGCTGTGGGGCATCTCGGGGACCGGACGAGATGCTGCGAGGGCGACACT 1080  
 DB 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401  
 QY 1081 GGGGGCCCATGCTGCTCTCTTCCACGCGACCTGCTTCTGGTGGCGCTGTGAGCTGG 1140  
 DB 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421  
 QY 1141 GGTGAGGCTGTGGGCTCTCTTCCAACTACGCGGCTTTACACCAAGTACGCGCTACCTTC 1200  
 DB 422 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTyrLeu 441  
 QY 1201 GACTGATCCATGGCGCATCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 DB 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456  
 RESULT 7  
 AAW25086  
 ID AAW25086 standard; Protein: 460 AA.  
 XX AC AAW25086;  
 XX DT 11-DEC-1997 (first entry)  
 XX DE Human protein C.  
 XX KW Protein C; transgenic animal; sheep; rabbit; cattle; goat; milk;  
 KW blood clotting; anticoagulant; human.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT 199..200  
 FT Cleavage-site /note= "two-chain cleavage site"  
 FT  
 PN W09720043-Al.  
 XX 05-JUN-1997.  
 XX 26-NOV-1996; 96WO-US18866.  
 XX 13-JUN-1996; 96US-0019692.  
 XX 30-NOV-1995; 95US-0565074.  
 XX (PPLT-) PPL THERAPEUTICS.  
 XX (ZYMO) ZYMOGENETICS INC.  
 XX Cottenham I, Foster DC, Garner I, Prunkard DE;  
 PI Sprecher CA, Temperley SM;  
 XX WPI: 1997-310599/28.  
 DR N-PSDB; AAT79723-24.  
 XX Production of protein C in transgenic animal - useful for high  
 quantity protein C production with therapeutic value  
 XX Disclosure; Page 58-60; 99pp; English.  
 XX This polypeptide comprises human protein C. A claimed method for  
 producing recombinant human protein C in the milk of a transgenic  
 animal involves: (a) providing a DNA construct comprising DNA  
 encoding a secretion signal and a protein C propeptide, operably  
 linked to DNA encoding two-chain cleavage site-modified protein C,  
 the 2 DNA sequences being linked to elements required for protein C  
 expression in a mammary gland of a host female animal; and (b)

CC using the DNA construct to breed a transgenic animal (esp. sheep,  
 CC rabbit, cattle, goat) that produces protein C in its milk, at  
 CC least 90% of the protein C being in the two-chain form.  
 CC Modification of the protein C two-chain cleavage site (see AAW25085)  
 CC improves the maturation of recombinant protein C from single chain  
 CC to two-chain form.  
 XX SQ Sequence 460 AA;

## Alignment Scores:

Pred. No.: 1.37e-159 Length: 460  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 Gaps: 0

S-09-763-153-1 (1-1245) x AAW25086 (1-460)

QY 1 GCCAACTCTCTCGAGGAGCTCCGTCACAGCAGCTGAGCGGAGTGCATAGAGAG 60  
 Db 43 AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 62  
 QY 61 ATCTGTGACTTCGAGAGGCCAGGAATTTTCCAAAATGTGGATGACACACTGGCCTC 120  
 Db 63 IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGTCTCAAGCAGCTGACGAGTGCACGAGTGTCTGGTCTGCTGGAGCAGCCGTCGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCTGTGTGGGCGGCGGCGCTTCGCCAGCGGAGTGCATGCGAGCTGCGACTGC 240  
 Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGGAGGCGGCTTCGCCAGCGGAGTGCATGCGAGCTGCGACTGC 300  
 Db 123 ArgSerGlyTrpGluArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACACGGGCGGCTGACCACTACTGCTAGAGGAGTGGGCTGGGCGGCTGAGCTCT 360  
 Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgGlySerCys 162  
 QY 361 GCGCTGTGCTACAGCTGGGAGGAGCTCTGCTGAGTGCATGCGAGCTGCGACTGC 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGGAGGCGCTGGAAGCGGATGGAGAGAGCGGAGTGCATGAGTGCATGAGTGCCT 480  
 Db 183 CysGlyArgProTyrLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATGTGAGGAGATGACGAGCGGGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGGAGGTGCTGCTGGACTCAAGAGAGAGTGGCTGGCGGCGGAGTGCCTC 600  
 Db 223 SerProTrpGlnValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGGGTGCTGACAGCGGCGGCTGAGTGCATGAGTGCATGAGTGCCT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaLysCysMetAspGluSerLysLysLeuLeu 262  
 QY 661 GTACGCTGGAGAGTATGACCTGGGCGGCTGGGAGAGTGGAGCTGGAGCTGGAGATC 720  
 Db 263 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGTCTTCGTCACCCCACTACAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGC 780  
 Db 283 LysGluValIlePheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGCAGCTGGCGGCGGCGGCGGCGGCTGCGAGAGCAGTGCCTGCTGCGGAG 840  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGCACAGCGCGAGCTCAATCAGCGCGGCCAGGAGACCTCGTGACGGGCTGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGTACACAGCAGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCGTCGTCGCCACCAATAGTCAGGAGGAGTGCATGAGCAACATGGTCT 1020  
 Db 363 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGTGGGCGGCTGCTGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTAGGCTGCTGGGCTGCTTCACAACTACGGGCTTTACACCAAGTACAGCGCTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGGATCCATGGGCGGACATCAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 443 AspIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## RESULT 8

AAP60001  
 ID AAP60001 standard; Protein; 461 AA.

AC AAP60001;

DF 25-JUL-1991 (first entry)

XX Sequence of polypeptide with human protein C activity.

DE Vascular disorder therapy; protein C deficiency.

KW Homo sapiens.

OS Location/Qualifiers

PH Key

FT Region

FT 1..32

FT /note="encoded by AAN60004"

FT 33..461

FT /note="encoded by AAN60001"

XX EP191606-A.

XX 20-AUG-1986.

XX 06-FEB-1986; 86EP-0300823.

XX 08-FEB-1985; 85US-0699967.

XX (ELIL) ELI LILLY & CO.

XX Bang NU, Beckmann RJ, Jaskunas SR, Lai MHT, Little SP;

PI Long GL, Santerre RF;

XX WPI; 1986-220077/34.

XX Prodn. of polypeptide having human protein C activity - is by

PT recombinant DNA procedures for prod. useful against vascular

FT disorders

XX Disclosure; Pages 10-12; 121pp; English.

XX The claimed sequence AAN60001 has "RIN-RW" attached to its 5' end

CC wherein: R= AAN60002 or AAN60003, and RI= AAN60004 or AAN60005; and M and

Tue Mar 18 16:19:30 2003

CC N= 0 or 1; provided that when M=0, N=0; and that when R= AAN60002, R1= AAN60004; and that when R= AAN60003, R1= AAN60005.

XX Sequence 461 AA;

Alignment Scores:  
Pred. No.: 1.37e-159 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 7 Gaps: 0

US-09-763-153-1 (1-1245) x AAP60001 (1-461)

QY 1 GCCAACTCTCTCTCGAGGAGTCGCTCACAGACGCTGGAGCGGGAGTCATAGAGGAG 60  
DB 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGGAGCCCAAGAAATTTCCAAAATGTGGATGACACACTGGCTTC 120  
DB 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82  
QY 121 TGTCTCAAGCACGTCGACGCTGACGAGTCTGTGGTCTTGGCTTGGAGCACCGTGGCC 180  
DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCTGTGCTTCGCGGACGCGCAGTCATCGACGCGCATCGGAGCTTCAGCTGCGACTGC 240  
DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGCGCTTCTGCCAGCGGAGGTGAGTCTCTCAATGTCTCGCTG 300  
DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACACGCGCTCCAGCGCTTACTGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTGT 360  
DB 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTCAAGCTGGGGAGGAGCTCTCTGAGTGTACCCCGAGTGAAGTTCCT 420  
DB 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGCCCTGGAAGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 183 CysGlyArgProTrpPlyArgGluGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACCGGGGAGAG 540  
DB 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGAGCTGTCTCTGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTACGCTTGGAGCTATGACCTGCGGCGCTGGGAGAGAGAGAGAGAGAGAGAGAG 720  
DB 263 ValArgLeuGlyGlyLysArgLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGGAGCTTCTCTCCACCCCACTACAGACAGACACACCGACATGACATGCGACTG 780  
DB 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGACCTGCGGACCGCCCGCCCTCTCGACAGCATAGTCCCATCTGCTCCCGGAC 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTGCAGAGCGGAGCTCAATCAAGCGCGGCGGAGGAGGAGGAGGAGGAGG 900

DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACCAGAGCCGAG 960  
DB 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAGATTCCCGTGTCCCGCACAAATGAGTGCAGCGAGGTCTATGAGCACACATGGTGTCT 1020  
DB 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGTGTGGGCGATCTCGGGGACCGGAGGATGCTCGGAGGCGGACAGT 1080  
DB 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGTGCTCTCTCCACGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGTCTCTTCAACTACGCTGCTTACACCAAGTACAGCGCTACCTC 1200  
DB 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
QY 1201 GACTGGATCCATGGACATCAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
DB 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
RESULT 9  
AAP70855  
ID AAP70855 standard; Protein: 461 AA.  
XX  
AC AAP70855;  
XX  
DT 10-MAY-1991 (first entry)  
XX Human Protein C.  
XX human Protein C; anti-coagulant; thrombosis; serine protease.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..42 /label= prepro leader peptide  
FT Disulfide-bond 59..64 /label= gamma-carboxyglutamic acid (Gla) domain  
FT Domain 60..63 /label= growth factor domains  
FT Disulfide-bond 92..111 /label= 183..319  
FT Disulfide-bond 101..106 /note= "links together the two processed chains"  
FT Disulfide-bond 105..120 /label= 373..387  
FT Disulfide-bond 122..131 /label= 398..426  
FT Disulfide-bond 140..151 /label= 238..254  
FT Disulfide-bond 147..160 /label= 139  
FT Disulfide-bond 162..175 /label= N-glycosylation site  
FT Domain 92..175 /label= N-glycosylation site  
FT Disulfide-bond 183..319 /label= N-glycosylation site  
FT Disulfide-bond 373..387 /label= N-glycosylation site  
FT Disulfide-bond 398..426 /label= N-glycosylation site  
FT Disulfide-bond 238..254 /label= N-glycosylation site  
FT Modified-site 139 /label= N-glycosylation site  
FT Modified-site 290 /label= N-glycosylation site  
FT Modified-site 355 /label= N-glycosylation site  
FT Modified-site 371 /label= N-glycosylation site  
FT Cleavage-site 211..212 /note= "in heavy chain; converts to activated protein C"  
FT Cleavage-site 197..198 /note= "apparent processing site for connecting"  
FT



FT Cleavage-site 199..200 dipeptide "  
 FT /note= "apparent processing site for connecting  
 FT dipeptide"  
 XX EP215548-A.  
 XX 25-MAR-1987.  
 XX 26-JUN-1986; 86EP-0304970.  
 XX 27-JUN-1985; 85US-0749600.  
 PR 15-AUG-1985; 85US-0766109.  
 XX (ZYMO-) ZYMOGENETICS INC.  
 PA (UNIW ) UNIV OF WASHINGTON.  
 XX Murray MJ, Berkner KL, Foster DC, Davie EW;  
 DR WPI; 1987-081505/12.  
 DR N-PSDB; AAN70102.  
 XX  
 PT Human protein C or activated protein C - prepd. using expression  
 PT vector capable of integration in mammalian host cell DNA  
 PS Claim 4; Fig 4; 52pp; English.  
 XX  
 CC Recombinantly produced protein C can be used to treat thrombotic  
 CC disorders such as venous thrombosis as it has anti-coagulant  
 CC properties. The protein sequence is thought to yield two peptide  
 CC chains; the first contains the Gla domain and growth factor domains  
 CC and the second (the activation peptide) contains the catalytic  
 CC domain.  
 XX  
 SQ Sequence 461 AA;

Alignment Scores:  
 Pred. No.: 1,37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 8 Gaps: 0

US-09-763-153-1 (1-1245) x AAP70855 (1-461)

1 GCCAACTCTCTCGAGAGCTCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluA-gGluCysIleGluGlu 62  
 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATCTGGATGACACACTGGCCTTC 120  
 63 IleCysAspPheGluGluAlaLysGluLeuPheGlnAsnValAspAspThrLeuAlaPhe 82  
 121 TGGTCCAAAGCAGCTGACCGTGACAGTGTGCTTGTGCTTGGAGCAGCGCTGGCC 180  
 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 181 AGCTCTGTCTGGGCGCAGCGACCTGATCGAGCGGATCGGAGCTTCAGCTCGACTGC 240  
 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 241 CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGCTTCCTCAATTGCTCGCTG 300  
 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 301 GACAAGCGGCTCGCAGCAGCTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgCysSerCys 162  
 361 CGCCTGGCTACAAGCTGGGGGAGCAGCTCTGCTGAGTGTACCCCGCAGTGAAGTCCCT 420  
 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182

QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGCAGTCACTGAAACGAGACACAGAA 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACAGCGCGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGCAGGTGGTCTCTGCTGACATCAAGAAGAGCTGGCTCGCGGCGAGTCTC 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
 QY 661 GTCAGGCTTGGAGAGTATGACTCGCGCGCTGGGAGAGTGGGAGCTGGAGCTGGAGATC 720  
 Db 263 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGGTCTTCGTCACCCCACTACAGCAAGAGACACCACCAATGACATCGCAGCTG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGACCTGGCCCGCCAGCGCCGCTCTCGCAGACCATAGTGCCTCTGCTCCCGGAC 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTTCAGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTCTGACGGCTGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCCGTGTCTCCGCAATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGTGGGCGCATCTCGGGGACCGGCGAGGATGCTGCGAGGCGGACAGT 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCTCTTCAACACTACGGCGTTTACACCAAGTACAGCGCTAGCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGATCCATGGGCATCAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 10  
 AAP90401  
 ID AAP90401 standard; protein; 461 AA.  
 XX  
 AC AAP90401;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Zymogen form of human protein C.  
 XX  
 KW Human protein C; zymogen form; activated C protein;  
 XX human liver mRNA; signal peptide; propeptide; antithrombotic.  
 OS Homo sapiens.  
 XX  
 PN EP323149-A.



AC AAR13074;  
 XX 02-OCT-1991 (first entry)  
 XX Protein C precursor.  
 XX Anticoagulant; fibrinolysis.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Peptide 2..42  
 XX Region 43..197  
 XX Cleavage-site /label= pre-pro peptide  
 XX Cleavage-site /label= light chain  
 XX Cleavage-site 197..198  
 XX Cleavage-site /label= proteolytic cleavage  
 XX Peptide 199..200  
 XX Peptide /label= proteolytic cleavage  
 XX Region 201..211  
 XX Region 212..461  
 XX Domain /label= activation peptide  
 XX Modified-site /label= heavy chain  
 XX Modified-site 43..79  
 XX Modified-site /label= Gla domain  
 XX Modified-site 48  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 49  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 56  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 58  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 61  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 62  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 67  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 68  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 71  
 XX Modified-site /label= gamma carboxyglutamic acid  
 XX Modified-site 139  
 XX Modified-site /label= N-glycosylation site  
 XX Modified-site 250  
 XX Modified-site /label= N-glycosylation site  
 XX Modified-site 355  
 XX Modified-site /label= N-glycosylation site  
 XX Modified-site 371  
 XX Modified-site /label= N-glycosylation site  
 XX WO9109951-A.  
 XX 11-JUL-1991.  
 XX 21-DEC-1990; 90WO-US07617.  
 XX 22-DEC-1989; 89US-0456092.  
 XX (TEIJ) TEIJIN LTD.  
 XX (ZYMO-) ZYMOGENETICS INC.  
 XX Foster DC, Holly RD, Suzuki M, Wakabayashi K, Kumar AA;  
 XX WPI; 1991-222903/30.  
 XX N-PSDB; AAQ12649.  
 XX Recombinant protein C with truncated light chain - for use as an  
 XX anticoagulant.  
 XX Disclosure; Fig 1; 60pp; English.

CC The sequence was deduced from a clone isolated from a cDNA library  
 CC prep'd. from mRNA from Hep G2 cells. It is a protein C precursor,  
 CC including light and heavy chains, which is cleaved to produce  
 CC activated protein C (see feature table). The DNA encoding the  
 CC sequence can be manipulated by genetic engineering techniques to  
 CC express a protein comprising (when activated) a heavy chain and a  
 CC truncated light chain comprising residues 1-149, 1-150, 1-151 or 1-  
 CC 152 of the natural sequence. The protein pref. comprises the  
 CC precursor of formula:  
 CC  
 CC pre-pro = pre-pro-L-X-H  
 CC The corresponding peptide of protein C with all/part replaced by  
 CC or prothrombin;  
 CC L = AAs 1-149, 150, 151 or 152 of light chain;  
 CC X = 3-10 Lys/arg residues; and  
 CC H = heavy chain.  
 CC Cells transformed with expression vectors contg. the modified DNA  
 CC sequences produce the new proteins which can be used to regulate  
 CC anticoagulant and fibrinolytic systems.  
 CC See also WO9112320 (AAR13074).  
 XX  
 XX Sequence 461 AA;  
 SQ

Alignment Scores:  
 Pred. No.: 1.37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 12 Gaps: 0

US-09-763-153-1 (1-1245) x AAR13074 (1-461)

QY 1 GCCAACTCTCTCTGGAGGAGCTCGGTACAGCAGCTGGAGGGAGTGCATAGAGGAG 60  
 DB 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 QY 61 ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAATGTGATGACACTGGCCTTC 120  
 DB 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGGTCAAGCAGCTCGAGCGGTGACCGTCTTGTCTTGCCTTGGAGCACCCTGGCGCC 180  
 DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCCTGTCTCGGGCACGCGCAGTGCATCGCGCATCGCGAGTTCAGTGGAGCTGC 240  
 DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGGAGGGCGCTTCTGCCAGCGCGAGGTGAGTCTTCAATTGCTCGCTG 300  
 DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAACGGCGCTGTCACGCAATTTACTGCTAGAGAGGTGGGCTGGCGCGCTGTAGCTGT 360  
 DB 143 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCCTGGCTACAGCTGGGGGACGACCTCTGCTGAGTGTACCCCGCAGTGAAGTTCCT 420  
 DB 163 AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCCCTTGAAGCGGATGGAGAGAGCGAGTACCTGAGAACGACAGACAGAA 480  
 DB 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGATGACACAGCGGGGAGAC 540  
 DB 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTTGGCAGGTGGTCTCTGCTGACTCAAGAGAGAGTGGCTGGCGGGGAGTGTCTC 600  
 DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242

601	ATCCACCCCTCCCTGGGTGCTGACAGCGGCCACTGCATGGATGAGTCCAGAGAGCTCCTT	660
243	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	262
661	GTCAAGGCTTGGAGATGATACCTCGCGGCTCGGAGAAAGTGGAGCTGGACCTCGACATC	720
263	ValArgLeuGlyGluIuTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	282
721	AAGGAGGCTTCGCTCCACCCCACTACAGACAGAGACACCGACCAATGACATCGCACTG	780
283	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu	302
781	CTGCACTTGGCCCGCGCGCACCTCTCCACACCATAGTGCCTCCCTCCCGGAC	840
303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	322
841	ACGGGCTTGGCAGCGCGAGCTCAATCAGCGCGGCCAGGAGACCTCGTGACGGCTGG	900
323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	342
901	GGCTACACAGCAGCGGAGAGAGAGGAGCCCAAGAGAAACCGCACCTTCGTCTCACTTC	960
343	GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	362
961	ATCAAGATCCCGTCCCGCACANTAGTGCAGCGAGGTGCATGAGCAACATGGTGCTCT	1020
363	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	382
1021	GAGNACATGCTGTGTCGGGCATCTCTCGGGGACCGCAGGATGCTCGGAGGCGACAGT	1080
383	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
1081	GGGGGGCCCATGTGCGCTTCCTTCACGCGCACCTGGTTCCTGGTGGGCTGGTGAGCTGG	1140
403	GlyGlyProMetValAlaSerPheHisGlyThrThrPheLeuValGlyLeuValSerTip	422
1141	GGTCAGGCTGTGGGCTTCCTTCAACTACGGGCTTTACACCAAGTACGCGGTACTCTC	1200
423	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	442
1201	GACTGGATPCCATGGGCACATCAGACACAGGAGCGCCCGCAGAAG	1245
443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 12		
AAIR13081	AAIR13081 standard; Protein; 461 AA.	
XX	AC	AAIR13081;
XX	DT	30-SEP-1991 (first entry)
XX	XX	Human protein C.
XX	XX	Phospholipid; binding protein; lipocortin; domain; vitamin K; PBP;
XX	XX	gla-domain; VKDP.
XX	OS	Homo sapiens.
XX	XX	Key
XX	XX	Location/Qualifiers
XX	XX	1..42
XX	XX	/label= sig_peptide
XX	XX	43..461
XX	XX	/label= mat_protein
XX	XX	W09105953-A.
XX	XX	11-JUL-1991.
XX	XX	13-DEC-1990; 90WO-US07335.
XX	XX	29-DEC-1989; 89US-0459082.
XX	XX	

PA	(ZYMO-) ZYMOGENETICS INC.
XX	Foster DC;
PPI	
XX	WPI; 1991-222905/30.
DR	N-PSDB; AAQ12678.
XX	Recombinant prodn. of hybrid phospholipid-binding proteins -
PPT	comprising lipocortin phospholipid-binding domain and
PT	vitamin K-dependent protein
XX	Disclosure; Fig 2; 57pp; English.
XX	This sequence, or a fragment of it, is used in the construction of
CC	hybrid phospholipid-binding proteins (PBP) having the same biological
CC	activity as human protein C or human activated protein C.
CC	The hybrid sequence would comprise at least one lipocortin phospholipid
CC	binding domain (PBD), e.g. of PAP-I, joined to a gla-domainless
CC	protein C or activated protein C. See AAQ12680-81 for such examples.
CC	See also AAQ12678-81.
XX	
SQ	Sequence 461 AA:
	Alignment Scores:
	Pred. No.: 1.37e-159 Length: 461
	Score: 2298.00 Matches: 415
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 98.33% Indels: 0
	DB: 12 Gaps: 0
	US-09-763-153-1 (1-1245) x AAR13081 (1-461)
Qy	1 GCCAACTCCTTCTTGAGGAGCTCGTCACAGCAGCTGGAGCGGGAGTGCATFAGAGGAG 60
Db	43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
Qy	61 ATCTGTCACTTCGAGGAGGCCAAGGAATAATTTCAAATGTGGATGACACACTGGCCTTC 120
Db	63 IleCysaspPheGluGluAlaLysGluIlePheGlnAsnValaspAspThrLeuAlaPhe 82
Qy	121 TGGTCCAAGCACGTCGACGGTGACCAGTCTGTGTCTTTGCCCTTTGGAGCACCCGTGGCC 180
Db	83 TrpSerLysHisValaspGlyaspGlnCysLeuValLeuProLeuGluHisProCysAla 102
Qy	181 AGCTGTGCTCGGGGACGGCAGCTGCATCATCGAGCGATCGGCGACTTCAGCTCGACTGC 240
Db	103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
Qy	241 CGCAGCGCTGGGAGGCGCGCTTCGCCAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300
Db	123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
Qy	301 GACAACGGCGCTGCACGCATTACTGTGCTAGAGAGGTTGGCTGGCGGCTGTAGCTGT 360
Db	143 AspasnGlyCysThrHisTyrCysLeuGluValglyTrpArgArgcysSerCys 162
Qy	361 GGCCCTGGCTACAAGCTGGGGACGACTCCTCCAGTGTGCACCCGACGTGAAGTTCCCT 420
Db	163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182
Qy	421 TGTGGAGGCCCTCGAAGCGGATGAGAAGCGCAGTACCTGAACACGACACACAGAA 480
Db	183 CysGLYArgProtrpLysArgMetGluLysLysArgSerHisLeuLysArgaspThrGlu 202
Qy	481 GACCAAGAAGACCAAGTAGATCCGCGTCAATTGATGGGAAGATGACACAGCGGGAGAC 540
Db	203 AspGlnGluaspGlnValaspProargLeulleaspGlyLysmethrArgargGlyasp 222
Qy	541 AGCCCCCTGGCAGTGTGCTCTGCTGACTCAAAGAAGAAGTGGCTGCGGGGCGAGTCTCT 600
Db	223 SerProTrpGlnValValleuleuaspSerLysLysLysLeuAlaCysGlyAlaValLeu 242



	FT	/label= C-terminal /note= "light chain" 200..211
	FT	/label= N-terminal
	FT	/note= "heavy chain"
	FT	458..461
	FT	/label= C-terminal
	FT	/note= "heavy chain"
	FT	451..461
	FT	/label= C-terminal
	FT	/note= "heavy chain"
	JPO5064588-A.	
	XX	
	PN	
	XX	
	PD	19-MAR-1993.
	PF	14-AUG-1991; 91JP-0228687.
	XX	
	PR	14-AUG-1991; 91JP-0228687.
	XX	(TEIJ ) TEIJIN LTD.
	PA	
	DR	WPI; 1993-128866/16.
	XX	
	XX	Human protein C and activated protein C with short H chains - useful as anti-clotting agents and fibrinolysis promoters
	PT	
	PS	Disclosure; Fig 1; 8pp; Japanese.
	CC	A human protein C or an activated protein C has a H chain contg. one of the residues 239-246 (= residues 450-457 in the sequence below) in the H chain of natural activated protein C as the C-terminal, or has a L chain contg. one of the residues 141-155 (= residues 141-155 in the sequence below), pref. residues 149-155 (= residues 149-155 in the sequence below) in the L chain of natural activated protein C as the C-terminal. The human protein C or the activated protein C can be used as an anticlotting agent or a fibrinolysis promoter.
	XX	Sequence 461 AA;
	SQ	
	Alignment Scores:	Length: 461
	Pred. No.:	1,37e-159 Matches: 415
	Score:	2298.00
	Percent Similarity:	100.00% Conservative: 0
	Best Local Similarity:	100.00% Mismatches: 0
	Query Match:	Indels: 0
	DB:	Gaps: 0
	US-09-763-153-1 (1-1245) x AAR34295 (1-461)	
	OY	1 GCCAACTCCTTCGTGGAGGACCTCCGTCACAGCGCTGGAGCGGAGTGCATAGAGGAG 60
	Db	43 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGlyCysIleGluGlu 62
	OY	61 ATCTGTGACTTCGAGGAGGCCAAGGAATTTTCCAAATGTGGATGACACACATGGCCTTC 120
	Db	63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82
	OY	121 TGGTCCAAGCAGCTCGACGGTGACCAGTGTGTGTTCCTTGGAGCACCGCTGCGCC 180
	Db	83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysSala 102
	OY	181 AGCTGTGCTGGGGCAGCGCACGTGCATCGACGGCATCGGCACGCTTCAGCTGCAGCTGC 240
	Db	103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
	OY	241 CGCAGCGGCTGGGAGGCCGCTTCTGCCAGCGGAGGTGAGCTTCCTCAATTGCTCCTG 300
	Db	123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
	OY	301 GACAACGGCGCTGCACGCATTACTCCCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360
	Db	143 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgArgCysSerCys 162
	OY	481 GACCAAGAAGACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACACAGCGGGGAGAC 540
	Db	203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp 222
	OY	541 AGCCCCCTGGCAGGTGGTCTCTGCTGACTCAAAGAAGAGTGGCCTGGCGGCGAGTGCNC 600
	Db	223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242
	OY	601 ATCCACCCCTCTGGTGCTTCACAGCGGCCACACGATGGATGATGATGATGATGATGATGAT 660
	Db	243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262
	OY	661 GTGAGCTGGAGAGTATGACCTGGCGGCTGGGAGAGTGGGAGCTGGACATC 720
	Db	263 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282
	OY	721 AAGGAGGTCTGTCACCCCACCTACAGCAGAGACACACCGACAATGACATCGCACTG 780
	Db	283 LysGluValPheValHisProasnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302
	OY	781 CTCGACCTGGCCAGCGCCGCTCTCCAGACCATAGTGGCCCATCTGCTCCCGGAC 840
	Db	303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322
	OY	841 AGCGCCCTTCAGAGCGCGAGCTCAATCAGCGCGCCGAGGAGACCTCGTACGCGGTGG 900
	Db	323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342
	OY	901 GGCTACACAGCAGCGGAGAGAGAGGCGCAAGAGAACCGACCTTGTCTCACTTC 960
	Db	343 GlyTrpHisSerSerArgGlyLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362
	OY	961 ATCAAGATTCCGTGGTCCCGCACAATGAGTGCAGCGAGGTGATGAGCAACATGTGCT 1020
	Db	363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382
	OY	1021 GAGAACATGCTGTGTGGGGCATCCTCGGGACCGGAGGATGCTCGGAGGGGACAGT 1080
	Db	383 GluAsnMetLeucysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402
	OY	1081 GGGGGGCCATGCTCGCTCTCCACGGCACCTGTTCTGTGGTGGCGCTGGTGGACTGG 1140
	Db	403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422
	OY	1141 GGTGAGGCTGTGGGCTCCTTCACAACTACGGCTTTACACCAAGTACGCGCTACCTC 1200
	Db	423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTyrrLeu 442
	OY	1201 GACTGGATCCATGGGCACATCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245
	Db	443 AsprilleHisGlyHisIleArgAspLysGluAlaProGlnLys 457
	SUULT 14	
	AAR34295	
	ID	AAR34295 standard; protein; 461 AA.
	XX	
	AC	AAR34295;
	XX	
	DT	10-AUG-1993 (first entry)
	XX	
	DE	Protein C.
	KX	protein C; heavy chain; light chain; anticomplementing; fibrinolysis;
	KW	promoter; anticoagulant.
	XX	
	OS	Homo sapiens.
	XX	Location/Qualifiers
	FH	Key
	FT	Peptide
	FT	/label= C-terminal
	FT	/note= "light chain"
	FT	193..197
	FT	Peptide



QY	361	GC	CC	TG	GG	TACA	AG	CTGG	GGG	GAC	CA	CTCT	CG	AG	TG	TCA	CC	CC	CG	CG	AG	TG	CC	CT	420	
Db	163	Ala	Pro	Gly	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	182	
QY	421	TG	GG	AG	CC	CT	GA	AG	CG	GG	AT	GG	AG	CG	GG	AT	GG	AG	CG	GG	AT	GG	AG	CG	480	
Db	183	Cys	Gly	Val	Pro	Trp	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	202	
QY	481	GAC	CA	AG	AG	CA	AG	AG	CA	AG	AG	CA	AG	AG	CA	AG	AG	CA	AG	AG	CA	AG	AG	CA	540	
Db	203	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	Val	Asp	Gln	222	
QY	541	AG	CC	CT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	600	
Db	223	Ser	Pro	Trp	Gln	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	660	
QY	601	AT	CC	AC	CC	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	720	
Db	243	Ile	His	Pro	Ser	Trp	Val	Leu	Thr	Ala	Ala	His	Cys	Met	Asp	Glu	Ser	Leu	Val	Leu	Val	Leu	Val	Leu	780	
QY	661	GT	CA	GG	CT	TG	AG	AT	GA	CT	CA	CT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	840	
Db	263	Val	Arg	Leu	Gly	Leu	Trp	Arg	Leu	Trp	Gly	Leu	Trp	Arg	Leu	Trp	Gly	Leu	Trp	Arg	Leu	Trp	Gly	Leu	900	
QY	721	AAG	AG	GT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	960	
Db	283	Lys	Glu	Val	Phe	Val	His	Pro	Asn	Trp	Ser	Lys	Ser	Trp	Thr	Asp	Asn	Phe	Val	Leu	Val	Leu	Val	Leu	1020	
QY	781	CT	CA	CT	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	CG	CC	1080	
Db	303	Leu	His	Leu	Ala	Gln	Pro	Ala	Thr	Leu	Ser	Gln	Thr	Leu	Val	Pro	Leu	Cys	Leu	Pro	Asp	322				
QY	841	AG	CG	CC	TT	CC	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1140	
Db	323	Ser	Gly	Leu	Ala	Glu	Arg	Glu	Leu	Asn	Gln	Ala	Gly	Gln	Leu	Thr	Leu	Val	Thr	Gly	Trp	342				
QY	901	GC	T	ACC	AC	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1200	
Db	343	Gly	Thr	His	Ser	Ser	Arg	Glu	Leu	Gly	Ala	Val	Arg	Asn	Arg	Thr	Phe	Val	Leu	Asn	Phe	362				
QY	961	AT	CA	AG	AT	CC	GT	GG	TCC	CC	CA	CA	AT	AG	TG	CA	AG	CG	CG	CG	CG	CG	CG	CG	1260	
Db	363	Ile	Val	Pro	Val	Val	Pro	His	Asn	Glu	Cys	Ser	Glu	Val	Met	Ser	Asn	Met	Val	Ser	382					
QY	1021	GAG	AA	CA	TG	CT	GT	CG	CG	CG	CA	CT	CT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1320	
Db	383	Glu	Asn	Met	Leu	Cys	Ala	Gly	Leu	Val	Leu	Gly	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	Leu	Val	1380	
QY	1081	GG	GG	GG	CC	CA	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	1440	
Db	403	Gly	Gly	Pro	Met	Val	Ala	Ser	Phe	His	Gly	Thr	Trp	Phe	Leu	Val	Gly	Leu	Val	Ser	Trp	422				
QY	1141	GG	T	GAG	GC	TG	CG	CT	CT	TCA	ACT	AC	GCG	CT	TTC	AC	CA	AG	TG	CT	TTC	AG	CT	TG	1500	
Db	423	Gly	Glu	Gly	Cys	Gly	Leu	Leu	His	Asn	Trp	Gly	Val	Thr	Lys	Val	Ser	Arg	Trp	Leu	442					
QY	1201	GAT	CG	AT	CC	AT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1560	
Db	443	Asp	Trp	Ile	His	Gly	Ile	Arg	Asp	Lys	Glu	Ala	Pro	Gln	Lys	578										
RESULT	15																									
AAW02600																										
ID	AAW02600	standard; Protein: 461 AA.																								
XX	XX																									
AC	AAW02600;																									
DT	05-NOV-1996	(first entry)																								
XX	Human protein C.																									
KW	Activated protein C; serine protease; thrombosis; thrombolytic; fibrinolytic; antithrombotic; blood clotting; therapy.																									

XX	OS	Homo sapiens.
XX	Key	Location/Qualifiers
FT	Peptide	/label= Pre-pro-peptide
FT	Protein	/label= Mat_protein
FT	Domain	/label= GLA_domain
FT	Disulfide-bond	/note= "forms disulphide bond with Cys111"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys105"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys101"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys120"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys92"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys106"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys131"
FT	Modified-site	/note= "forms disulphide bond with Cys122"
FT	Disulfide-bond	/label= N-glycosylation_site
FT	Disulfide-bond	/note= "forms disulphide bond with Cys151"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys160"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys140"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys147"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys175"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys162"
FT	Misc-difference	/note= "forms disulphide bond with Cys319"
FT	Misc-difference	/note= "residue 196 is replaced by Lys, Arg or in constructs of the invention"
FT	Misc-difference	/note= "residues 198-199 are replaced by Lys-Lys or Arg-Arg in constructs of the invention"
FT	Misc-difference	/note= "residue 200 is replaced by Ala, Ser, Thr or Gly in constructs of the invention"
FT	Cleavage-site	/note= "cleavage site for connecting dipeptide"
FT	Cleavage-site	/note= "cleavage site between connecting dipeptide and activation peptide"
FT	Peptide	200..211
FT	Cleavage-site	/label= Activated_protein-C
FT	Disulfide-bond	/note= "cleavage site for activation peptide"
FT	Disulfide-bond	/note= "forms disulphide bond with Cys254"
FT	Modified-site	/note= "forms disulphide bond with Cys238"
FT	Disulfide-bond	/label= N-glycosylation_site
FT	Modified-site	/note= "forms disulphide bond with Cys183"
FT	Modified-site	/label= N-glycosylation_site
FT	Modified-site	371

Tue Mar 18 16:19:30 2003

FT Disulfide-bond 373 /label= N-glycosylation\_site  
 FT 387 /note= "forms disulphide bond with Cys387"  
 FT Disulfide-bond 387 /note= "forms disulphide bond with Cys373"  
 FT 398 /note= "forms disulphide bond with Cys426"  
 FT Disulfide-bond 426 /note= "forms disulphide bond with Cys398"  
 FT 426 /note= "forms disulphide bond with Cys398"

XX US5516650-A.  
 XX 14-MAY-1996.  
 XX 27-JUN-1985; 85US-0749600.  
 XX 28-FEB-1989; 89US-0317205.  
 XX 27-JUN-1985; 85US-0749600.  
 XX 29-OCT-1986; 86US-0924462.  
 XX 08-DEC-1987; 87US-0130370.  
 XX 10-SEP-1990; 90US-0582131.  
 XX 04-DEC-1992; 92US-0987532.  
 XX 08-APR-1994; 94US-0225253.  
 XX (ZYMO ) ZYMOGENETICS INC.

XX Berkner KL, Foster DC, Murray MJ;  
 XX WPI; 1996-251006/25.  
 XX N-PSDB; AAT32795;  
 XX N-PSDB; AAT32796.

XX New DNA encoding modified forms of opt. activated protein C - and  
 XX related transformed cells for prodn. of recombinant protein C for  
 XX use e.g. as an anti-thrombotic agent

XX Example 1; Fig 2A-C; 34pp; English.

XX Human protein C (AAW02600) is a zymogen of a serine protease that  
 XX plays an important role in the regulation of blood coagulation  
 XX and the generation of fibrinolytic activity in vivo. It is  
 XX synthesised in the liver and processed to a 2-chain molecule,  
 XX which is itself converted to activated protein C. Protein C and  
 XX activated protein C are useful in the treatment of thrombotic  
 XX disorders. They can be produced e.g. in mammalian host cells using  
 XX a cDNA clone (AAT32795) derived from Hep G2 cells. Variant protein  
 XX C, modified to improve cleavage between the heavy and light  
 XX chains of the circulating intermediate, can also be produced.

XX Sequence 461 AA;  
 XX Segment Scores:  
 XX Id. No.: 1.37e-159 Length: 461  
 XX Score: 2298.00 Matches: 415  
 XX Percent Similarity: 100.00% Conservative: 0  
 XX Best Local Similarity: 100.00% Mismatches: 0  
 XX Query Match: 98.33% Indels: 0  
 XX DB: 17 Gaps: 0

US-09-763-153-1 (1-1245) x AA02600 (1-461)

QY 1 GCCAACTCTCTGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 QY 61 ATCTGTGACTTCGAGGAGCCCAAGGAAATTTTCCAAAATGTGGATGACACTGGCCCTTC 120  
 Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGGTCCAAAGCAGCTCAGCGGTGACACAGTCTTGGCTTGGCTTGGAGCAGCCCGTGGCC 180  
 Db 83 TrpSerLysHisValAspGlyGlnCysLeuValLeuProLeuGluHisProCysAla 102

QY 181 AGCTGTGCTGGCGGACGGCAGCTGCATGCAGCGCAGCTTCAGTGCAGTGCAGTGC 240  
 Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGAGGCGCTTCTCCAGCGCGAGGTGAGCTTCTCAATGTCTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAACGGCGCTGCACGCAATTAATCTAGAGAGGTGGCTGGCGGCTCTAGCTGT 360  
 Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCTGCTACAAAGCTGGGGACGACCTCTCAGTGTCAACCCCGCAGTGAAGTCCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTGGAAGCGATGGAGAGCGAGTCCACCTGAAAGGAGACAGACAGAA 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCAATTGATGGAGATGACCGCGGCGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp 222  
 QY 541 AGCCCTGCGCAGGTGCTCTGCTGACTCAAGAGAGAGTGGCTGGCGGCGGCGAGTCTC 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCTCTCTGGTGTGTGACAGCGGCCACTGCATGGATGAGTCCAGAGAGTCTCT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 262  
 QY 661 GTCAGGCTTGGAGTATGACCTGGCGCTGGGAGAGTGGAGTGGAGTGGAGTGGAGTGC 720  
 Db 263 ValArgLeuGlyGlyTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGTCTTCTGCTCCACCCCAACTACAGAGAGAGACACCCAGCAATGACATGCAC 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 302  
 QY 781 GTGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACACAGCAGCGGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCCGTGGTCCCGCACAAATGATGTCAGCGGAGTGCATGAGCAACATGGTCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATCTGTGTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 Db 403 GlyClyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCGTCTGGGCTCTTCAAACTACGGCGGTTTACACAAAGTACGGCGGCTACCTC 1200  
 Db 423 GlyGlyGlyCysGlyLeuLeuHisAsnTyrGlyValThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGTGATCCATGGCGCATCATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
 RESULT 16

AAV49561  
 ID AAY49561 standard; Protein; 461 AA.  
 XX AC AAY49561;  
 XX DT 13-JAN-2000 (first entry)  
 XX DE Human lecithin cholesterol acyltransferase protein sequence.  
 XX KW Human; coding sequence polymorphism; vascular pathology gene;  
 KW polymorphic site; phenotype correlation; forensic; paternity testing;  
 KW medicine; genetic analysis; vascular disease.  
 OS Homo sapiens.  
 XX PN W09950454-A2.  
 XX DT 07-OCT-1999.  
 XX PF 26-MAR-1999; 99WO-US06473.  
 XX PR 01-APR-1998; 98US-0054272.  
 XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.  
 XX PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;  
 XX DR WPI: 1999-620066/53.  
 XX DR N-PSDB; AAZ32180.  
 XX PT Determination of polymorphisms in genes, especially those identifying  
 PT predisposition to vascular disease -  
 XX PS Disclosure; Fig 24; 134pp; English.  
 XX CC AAZ32159 to AAZ32194 represent reference alleles for specifically  
 CC claimed nucleic acid sequences from the present invention which comprise  
 CC polymorphic sites as given in a table in the specification, selected  
 CC from 92 single nucleotide polymorphisms in which the nucleotide at the  
 CC polymorphic site is different from a nucleotide at the same site in a  
 CC reference allele. The nucleic acids, and primers and probes, are used to  
 CC identify polymorphisms, which may predispose an individual to disease,  
 CC especially a vascular disease. They can also be used in phenotype  
 CC correlations, forensics, paternity testing, medicine or genetic  
 CC analysis. AAY49550 to AAY49573 represent the proteins which correspond  
 CC to some of the reference alleles.

Sequence 461 AA;  
 Alignment Scores:  
 Pred. No.: 1.37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 20 Gaps: 0

US-09-763-153-1 (1-1245) x AAY49561 (1-461)  
 QY 1 GCCACTCTCTCGTGGAGGAGCTCGTCCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluLucyileGluGlu 62  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 Db 63 IleCysAspPheGluGluAlaLysGluilePheGlnAsnValAspAspPheLeuAlaPhe 82  
 QY 121 TGGTCCAGCAGCTGCAGCGGTGACAGTGTGCTTGGCTTGGCCCTTGAGCACCCTGGCGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGluCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCCTGTCTCGGGGACGGCAGCTGCATCGACGATCGGAGCTTCAGCTGCCACTGCC 240  
 Db 181 AGCCTGTCTCGGGGACGGCAGCTGCATCGACGATCGGAGCTTCAGCTGCCACTGCC 240

Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGGCTGGGAGGCGCGCTTCGCCAGCGCGAGGTGAGCTTCCTCAATGCTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAACGGCGGTGCACGCTTACTGCTAGAGGAGGTGGCTGGCGCGCTGAGCTGT 360  
 Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GGCCTTGGCTACAAGCTGGGCGAGCAGCTCTGAGTGTACCCCGCAGTGAATTCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTTGAAGCGGATGAGAGGAGCGGCTCATTGATGGGAGATGACAGGCGGAGAC 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTGATGGGAGATGACAGGCGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGCAGGTGCTCTGCTGACTCAAGAAGAAGCTGGCTGGCGGCGAGTGTCT 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGGTGGTGTGACAGCGGCGCTGATGATGATGATGATGATGATGATGAT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 262  
 QY 661 GTCAGGCTGGAGAGTATGACCTGGCGGCTGGGAGAGTGGGAGCTGGAGCTGGAGCATC 720  
 Db 263 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspLeu 282  
 QY 721 AAGAGGTCTTCTGTCACCCCACTACAGAACAGCACCACCAATGACATGCGACTG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspLeuAlaLeu 302  
 QY 781 CTGCACTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GCTACACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCCGTGGTCCCGCACAATGAGTGCAGGAGGTGATGAGCAACATGCTGTCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGCGGCGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCTTACAACTACGCGGCTTTACCAAAAGTCAGCGCGCTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGATCCATGGGCATCATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 17  
 AAE08626  
 ID AAE08626 standard; Protein; 461 AA.

Tue Mar 18 16:19:30 2003

XX AC AAE08626;  
 XX 01-NOV-2001 (first entry)  
 XX DE Human wild type protein C.  
 XX  
 XX Human: protein C derivative; anticoagulation activity; thrombosis;  
 KW serpin inactivation; acute coronary syndrome; myocardial infarction;  
 KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;  
 KW disseminated intravascular coagulation; DIC; burn; transplantation;  
 KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;  
 KW haemolytic uremic syndrome; acute arterial thrombotic occlusion;  
 KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1-42  
 XX Protein /label= Signal\_peptide  
 XX /label= Mature\_human\_wild type\_protein\_C  
 XX  
 XX WO200159084-A1.  
 XX 16-AUG-2001.  
 XX 02-FEB-2001; 2001WO-US01221.  
 XX 11-FEB-2000; 2000US-0181948.  
 XX 14-MAR-2000; 2000US-0189199.  
 XX (ELIL ) LILLY & CO ELI.  
 XX Gerlitz BE, Grinnell BW, Jones BE;  
 XX WPI; 2001-514662/56.  
 XX N-PSDB; AAD15224.  
 XX  
 XX Protein C derivative for treating acute coronary syndromes, vascular  
 XX occlusive disorders, thrombotic disorders and sepsis, comprises  
 XX substitutions at specified amino acid positions -  
 XX  
 XX Disclosure; Page 44-46; 59pp; English.  
 XX  
 XX The invention relates to human protein C derivatives and nucleic acid  
 XX molecules encoding such derivatives. These derivatives have increased  
 XX anticoagulation activity, resistance to serpin inactivation and  
 XX increased sensitivity to thrombin activation compared to wild type  
 XX protein C, and retains the biological activity of the wild type human  
 XX protein C, and protein C derivatives are useful in the manufacture of a  
 XX medicament for the treatment of acute coronary syndromes e.g. myocardial  
 XX infarction and unstable angina; and disease states predisposing to  
 XX thrombosis; vascular occlusive disorders and hypercoagulable states e.g.  
 XX disseminated intravascular coagulation (DIC), burns, transplantations,  
 XX thalassaemia, sickle cell disease, viral haemorrhagic fever and  
 XX haemolytic uremic syndrome; sepsis in combination with bacterial  
 XX permeability increasing protein; thrombotic disorders in combination  
 XX with an anti-platelet agent; protein C deficiency; acute arterial  
 XX thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral  
 XX or peripheral arteries or in vascular grafts in combination with a  
 XX thrombolytic agent. Nucleic acid molecules of the invention are useful  
 XX for treating humans with genetically predisposed prothrombotic disorders  
 XX by gene therapy. The present sequence is human wild type protein C.  
 XX  
 XX SQ Sequence 461 AA;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 1.37e-159 Length: 461  
 XX Score: 2298.00 Matches: 415  
 XX Percent Similarity: 100.00% Conservative: 0  
 XX Best Local Similarity: 100.00% Mismatches: 0  
 XX Query Match: 98.33% Indels: 0

DB: 22 Gaps: 0  
 US-09-763-153-1 (1-1245) x AAE08626 (1-461)  
 QY 1 GCCAACTCTCTCGAGGAGCTCCGTACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluGluGlu 62  
 QY 61 ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCTTC 120  
 Db 63 IleCysAspPheGluGluAlaLysGluLeuPheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGTCTCAACGACGTCGACGCTGACAGCTGCTGTGTGCTTCCCTTGGAGCAGCTGGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGlyAspGlyLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCTGTGTGGGGCAGCGCAGCTGCATCGACGCGCATCGCAGCTCAGCTGGAGTGC 240  
 Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGAGGTGAGCTTCTCCTCAATTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAACGGCGCTGCACGCAATTACTGCTAGAGGAGGTGGCTGGCGGCTGTAGCTGT 360  
 Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCTGCTTACAAAGCTGGGGGAGCGCTCTCTCAGTGTACCCCGCAGTGAAGTTCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspLeuGluGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTCGAGCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATGATGGGAGATGACCGAGGCGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
 QY 661 GTCAGGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGGAGCTGGAGCTGGAGATC 720  
 Db 263 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAlaLeuLeu 302  
 QY 781 CTGACCTGCGCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGGAGAGCGCGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGlnGlnGlnGlnGlnGln 342  
 QY 901 GGTACACACAGCGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluLysLysLysLysLysLysLysLysLysLys 362  
 QY 961 ATCAAGATTCCCGTGGTCCCGCAACAATAGTCAGCGAGGAGTGCATGAGCAACATGTTCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGTGGGGGATCTCTCGGGGACCGCGGAGGATGCTCGGAGGCGGACGT 1080

Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGCCCATGGCTCCTTCCACGGCACCTGGTTCCTGGTGGCTGGTGGCTGG 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCCCTCACAACACTACGCGGTTTACACCAAGTCAGCGGTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 442  
 QY 1201 GACTGGATCCATGGGCACATCAGACAGCAAGGAAGCCGCCAGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## RESULT 18

B82674

AA82674 standard; Protein; 461 AA.

AC AAB82674;

DT 15-OCT-2001 (first entry)

DE Wild-type human protein C.

XX Protein C; human; coronary syndrome; thrombosis; angina;

KW myocardial infarction; vascular occlusive disorder;

KW hypercoagulation; sepsis; protein C deficiency; occlusion;

KW thromboembolism; stenosis; antibacterial; immunosuppressive;

KW thrombolytic; cardiac; antidiabetic; antidiabetic; therapy.

OS Homo sapiens.

XX Key

FT Location/Qualifiers

FT 1..42

FT /label= Signal\_peptide

FT 43..461

FT /label= Mature\_protein

FT 43..87

FT /note= "Gla domain"

FT Disulfide-bond

FT 92..111

FT Disulfide-bond

FT 101..106

FT Disulfide-bond

FT 122..131

FT Disulfide-bond

FT 140..151

FT Disulfide-bond

FT 162..175

FT Disulfide-bond

FT 183..319

FT Disulfide-bond

FT 238..254

FT Disulfide-bond

FT 373..387

FT Disulfide-bond

FT 398..426

FT Cleavage-site

FT 198..199

FT /note= "cleavage makes a 2-chain inactive

FT precursor (155-amino acid light chain

FT attached via a disulfide bond to a

FT 262-amino acid heavy chain)"

FT Modified-site

FT 48

FT /note= "gamma-carboxylated"

FT Modified-site

FT 49

FT /note= "gamma-carboxylated"

FT Modified-site

FT 56

FT /note= "gamma-carboxylated"

FT Modified-site

FT 58

FT /note= "gamma-carboxylated"

FT Modified-site

FT 61

FT /note= "gamma-carboxylated"

FT Modified-site

FT 62

FT /note= "gamma-carboxylated"

FT Modified-site

FT 67

FT /note= "gamma-carboxylated"

FT Modified-site

FT 68

FT /note= "gamma-carboxylated"

FT Modified-site

FT 200..211

FT /note= "activation peptide; removal activates the

FT 2-chain zymogen"

FT Cleavage-site 211..212  
 FT /note= "thrombin cleavage site"  
 FT Modified-site 139  
 FT /note= "N-glycosylated"  
 FT Modified-site 290  
 FT /note= "N-glycosylated"  
 FT Modified-site 355  
 FT /note= "N-glycosylated"  
 FT Modified-site 371  
 FT /note= "N-glycosylated"  
 FT Modified-site 44  
 FT /note= "O-phosphorylated"  
 XX  
 PN WO200157193-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 19-JAN-2001; 2001WO-US00020.  
 XX  
 XX 02-FEB-2000; 2000US-0179801.  
 PR 14-MAR-2000; 2000US-0189197.  
 XX  
 XX (ELIL ) LILLY & CO ELI.  
 XX  
 XX Gerlitz BE, Jones BE;  
 XX  
 XX WPT; 2001-496919/54.  
 DR N-PSDB; AAB26362.  
 XX

Novel human protein C derivative for treating, e.g., myocardial infarction, unstable angina, sepsis, thrombotic disorders, acute arterial thrombotic occlusion, and thromboembolism - Disclosure; Page 50-52; 63pp; English.

The present sequence is that of human protein C prepro-polypeptide. The invention relates to human protein C derivatives having at least 2 amino acid substitutions, and to recombinant DNA molecules encoding such derivatives. These derivatives have increased anticoagulant activity and resistance to inactivation by serpins compared with wild-type human protein C but retain the biological activity of the wild-type protein. The amino acid substitutions are selected from H10Q, S11G, S12K, Q32E, N33D, N33F, and amino acids at positions 194, 195, 228, 249, 254, 302, or 316 of the mature protein C. Polypeptide substituted with Ser, Ala, Thr, His, Lys, Leu, Arg, Asn, Asp, Glu, Gly or Gln (numbering relative to the protein C mature protein sequence). Preferred protein C derivatives are given in AAB82675-78. Also claimed are a vector comprising DNA encoding the novel human protein C derivatives, transformed host cells and a method of producing the human protein C derivatives. The protein C derivatives are useful for treating coronary syndromes and disease states predisposing to thrombotic disorders. myocardial infarction and unstable angina), vascular occlusive disorders and hypercoagulable states, sepsis (in combination with bactericidal permeability increasing protein or with tissue factor pathway inhibitor), thrombotic disorders (in combination with an anti-platelet agent or by local delivery through an intracoronary catheter), protein C deficiency, acute arterial thrombotic occlusion, thromboembolism, or stenosis in coronary, cerebral or peripheral arteries or in vascular grafts. Human patients with genetically predisposed thrombotic disorders may be treated by gene therapy (all claimed).

SQ Sequence 461 AA;

## Alignment Scores:

Pred. No.: 1 37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 22 Gaps: 0

us-09-763-153-1-n2p.rag

Tue Mar 18 16:19:30 2003

QY	1081	GGGGGGCCCATGGTGGCTCTTCCAGCGCACCTGGTCTCTGGTGGGCTGGTGGAGCTGG	1140
Db	403	GlyGlyPrometValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	422
QY	1141	GGTGAGGCTGGGGCTCTTCCAGCACTAGCGGCTTTACACCAAGTACCGCTACCTC	1200
Db	423	GlyGlyGlyGlyGlyGlyLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	442
QY	1201	GACTGGATCCATGGGACATCAGACAGCAAGCAAGCCGCCAGAG	1245
Db	443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 19			
ID	AAB36895	standard; Protein; 461 AA.	
XX	AC	AAB36895;	
XX	XX	26-FEB-2001 (first entry)	
XX	DE	Human protein C derivative 2.	
XX	KW	Protein C; human; vascular occlusive; burn; transplantation;	
XX	KW	deep vein thrombosis; sickle cell; thalassemia;	
XX	KW	thrombotic disorders; myocardial infarction; angina; stroke.	
XX	OS	Homo sapiens.	
XX	XX	W0200066754-AL.	
XX	PD	09-NOV-2000.	
XX	XX	13-APR-2000; 2000WO-US08722.	
XX	PF	30-APR-1999; 99US-0131801.	
XX	PR	(ELIL) LILLY & CO ELI.	
XX	PA	Gerlitz BE, Jones BE;	
XX	PI	WPI; 2001-007227/01.	
XX	DR	N-PSDB; AAC83312.	
XX	XX	protein C derivatives, useful for treating vascular occlusive disorder,	
XX	PT	hypercoagulable state, thrombotic disorder and disease states	
XX	PT	predisposing thrombosis, comprises specific amino acid substitutions -	
XX	PS	Claim 1; Page 44-46; 57pp; English.	
XX	CC	The present invention relates to a human protein C derivative. The	
XX	CC	protein is useful for treating vascular occlusive disorders, such as	
XX	CC	hypercoagulable states such as sepsis, disseminated intravascular	
XX	CC	coagulation, purpura fulminans, major trauma, major surgery, burns,	
XX	CC	adult respiratory distress syndrome, transplantation, deep vein	
XX	CC	thrombosis, heparin-induced thrombocytopenia, sickle cell disease,	
XX	CC	thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic	
XX	CC	purpura, and hemolytic uremic syndrome, and also useful for treating	
XX	CC	infarction, unstable angina, and stroke. Protein C derivatives with	
XX	CC	amino acid substitutions result in increased resistance to	
XX	CC	inactivation by serpins when compared to wild-type activated human	
XX	CC	protein C. They also have longer half-lives in human blood and hence	
XX	CC	require either less frequent administration and/or smaller dosage	
XX	CC	than wild type human protein C for treating disorders.	
XX	SQ	Sequence 461 AA;	
XX	XX	Alignment Scores: 1.37e-159 Length: 461	
XX	XX	Pred. No.: 2298.00 Matches: 415	
XX	XX	Score: 100.00% Conservative: 0	
XX	XX	Percent Similarity: 100.00% Mismatches: 0	
XX	XX	Percent Local Similarity: 100.00%	



Query Match:	98.33%	Indels:	0
DB:	22	Gaps:	0
US-09-763-153-1 (1-1245) x AAB36895 (1-461)			
QY	1	GCACACCTCTCTGGAGGAGCTCCGTCACACGAGCGGAGTGCATAGAGGAG	60
Db	43	AlaAsnSerPheLeuGluLeuArgHisSerLeuGluArgGluCysIleGluGlu	62
QY	61	ATCTGTGACATCCAGAGGCCAAGGAAATTTCCAAATATGAGTACACACTGGCTTC	120
Db	63	IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
QY	121	TGTTCAAGACCTGCACGGTACCAGTGTCTGCTGGCCCTGGAGCACCGCTGC	180
Db	83	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
QY	181	AGCTGTGCTGGGCGACGACCTGATCGAGCGGATCGGAGTGGCTGGCGCTGAGCTG	240
Db	103	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
QY	241	CGCAGCGCTGGAGGCGCTCTGTCAGCGGAGTGGCTGGCGCTGAGCTGAGCTG	300
Db	123	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
QY	301	GACACGGCGCTGCACGCTACTCTGCTAGAGAGTGGCTGGCGCTGAGCTG	360
Db	143	AspAsnGlyGlyCysThrHisGlyCysLeuGluGluValGlyTrpArgCysSerCys	162
QY	361	CGCGCTGCTACAGCTGGGAGCAGCTCTGCTGAGTGCACCGCGAGTGAAGTTCCT	420
Db	163	AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro	182
QY	421	TGTTGGAGGCGCTGGAAGCGATGAGAGAGCGAGTCACTGAAACGAGACACAGAA	480
Db	183	CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu	202
QY	481	GACCAAGAACCAAGTATCCGCGCTCATTTGATGGAGAGTACACCGCGGGAGAC	540
Db	203	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	222
QY	541	AGCCCTGGAGTGGCTGCTGCTGAGTCAAGAGAGCTGCTGGCGGAGTGGCTG	600
Db	223	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	242
QY	601	ATCCACCCCTCTGGGTCGACAGCGGCGCCACTGATGATGATGATGATGATGATG	262
Db	243	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	282
QY	661	GTCAGGCTGGAGATGACCTGGCGGCTGGGAGTGGAGTGGAGTGGAGTGGAGTGG	302
Db	263	ValArgLeuGlyLysTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle	322
QY	721	AAGGAGTCTGCTCCACCCCACTACAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAG	342
Db	283	LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu	362
QY	781	CTGCACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	382
Db	303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	402
QY	841	AGCGGCTTGCAGAGCGGAGCTCAATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG	422
Db	323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	442
QY	901	GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	462
Db	343	GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	482
QY	961	ATCAAGATTCCTGTCCTCCGCAATGATGTCGAGCGGAGTCAAGACATGGTGTCT	502
Db	363	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	522
QY	1021	GAGAACATGCTGTGTCGGGCAATCTCTGGGAGCGGAGTGGCTGGAGCGGAGTGG	1080
Db	383	GluAsnMetLeuGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
QY	1081	GGGGGGCCCATGTCCTCCCTCCCTCCACGCGACCTGGTTCCTGGTGGGCGCTGG	1140
Db	403	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	422
QY	1141	GTTGAGGCTGGGCTCTTCACACCTACCTACCTACCTACCTACCTACCTACCTACCT	1200
Db	423	GlyGlyGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTrpLeu	442
QY	1201	GACTGCATCCATGGGCGACATCAGACACAGGAGGAGGAGGAGGAGGAGGAGGAGG	1245
Db	443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 20			
AAU99001			
ID	AAU99001	standard; Protein; 461 AA.	
XX	AAU99001;		
AC	AAU99001;		
DT	23-AUG-2002	(first entry)	
DE	Human Protein C precursor protein.		
KW	Human; Protein C; N-glycosylation; APC; activated protein C; precursor;		
KW	serum half-life; chromosome 2q13-q14; stroke; myocardial infarction;		
KW	sepsis; septic shock; disseminated intravascular coagulation; DIC;		
KW	bone marrow transplantation; major surgery; trauma; ARDS; coagulant;		
OS	Homo sapiens.		
Key	Location/Qualifiers		
FT	1..42		
FT	/label= Signal_peptide		
FT	43..461		
FT	/label= Mature_protein_C		
FT	43..197		
FT	/label= Light_chain		
FT	198..199		
FT	/label= Lys_Arg_dipeptide		
FT	200..301		
FT	/label= Activation_peptide		
FT	200..461		
FT	/label= Heavy_chain		
PN	WO200232461-A2.		
XX			
XX	25-APR-2002.		
XX	15-OCT-2001; 2001WO-DK00679.		
XX	18-OCT-2000; 2000DK-0001560.		
XX	18-OCT-2000; 2000US-242268P.		
XX	21-JUN-2001; 2001DK-0000970.		
XX	21-JUN-2001; 2001US-300154P.		
XX	(MAXY-) MAXYGEN APS.		
XX	(MAXY-) MAXYGEN HOLDINGS LTD.		
XX	Andersen KV, Pedersen AH, Freskgaard PO;		
XX	WPI; 2002-489875/52.		
XX	N-PSDB; ABR86038.		
XX	Novel conjugate useful for treating or preventing septic shock, stroke		
XX	and myocardial infarction, comprises non-polypeptide group covalently		
XX	attached to protein C polypeptide comprising an attachment group		
XX	Example 4; Page 76-77; 92pp; English.		

sequence 461 AA;

152-1 / 1-1245) X AAU99001 (1-461)

17 2003. 17:31:03

Search completed: M  
Job time : 83 secs

[illegible]

17 2003. 17:31:03

Search completed: M  
Job time : 83 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:26:59 ; Search time 51 Seconds

(without alignments)  
4693.622 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 263224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p\_model -DEV=xlp  
-O=/cgn2.1/USPTO\_spool/US09763153/runat\_11032003\_084248\_2005/app\_query.fasta\_1.1415  
-DB=PIR\_73 -QFMT=fastcan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=60  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153.scgn.1.1.38 @runat\_11032003\_084248\_2005 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	461	1 KXHU	protein C (activat
2	1660	71.0	456	1 KXBO	protein C (activat
3	1650.5	70.6	461	1 S18994	protein C (activat
4	1637.5	70.1	461	1 JX0210	protein C (activat
5	818.5	35.0	482	1 EXRT	coagulation factor
6	809	34.6	482	1 EXBO	coagulation factor
7	801.5	34.3	475	1 EXCH	coagulation factor
8	801.5	34.3	488	1 EXHU	coagulation factor
9	783	33.5	466	1 KFHU7	coagulation factor
10	779.5	33.4	407	1 KFB07	coagulation factor
11	769.5	32.9	43	2 I46332	coagulation factor
12	763	32.6	432	1 A30351	coagulation factor
13	736	31.5	461	1 KFHU	coagulation factor
14	726	31.1	459	2 JQ0419	coagulation factor

15	714.5	30.6	416	1 KFBO	coagulation factor
16	562.5	24.1	622	1 TBHU	thrombin (EC 3.4.2
17	538.5	23.0	625	1 TBBO	thrombin (EC 3.4.2
18	533	22.8	618	2 A35827	thrombin (EC 3.4.2
19	525.5	22.5	617	2 S10511	thrombin (EC 3.4.2
20	470	20.1	655	1 A46688	hepatocyte growth
21	468	20.0	422	1 KXHU2	plasma protein z p
22	445.5	19.1	396	1 KXBOZ	plasma protein z -
23	427	18.3	271	2 I46580	factor IX - pig (f
24	426	18.2	638	1 KQMSPL	plasma kallikrein
25	424.5	18.2	638	1 KQHUP	plasma kallikrein
26	420.5	18.0	699	1 I54763	Ra-reactive factor
27	420	18.0	27	2 I46712	factor IX - rabbit
28	419.5	18.0	812	1 PLMS	plasma (EC 3.4.21
29	419	17.9	625	1 KFHU1	coagulation factor
30	418.5	17.9	855	2 JC7731	membrane-bound arg
31	418.5	17.9	855	2 JC7775	membrane type-sei
32	417	17.8	285	2 I48144	coagulation factor
33	414	17.7	638	1 KQRTPL	plasma kallikrein
34	411	17.6	786	1 A47547	serine proteinase
35	410.5	17.6	275	2 C35863	trypsin (EC 3.4.2
36	410	17.5	560	1 JC4795	plasma hyaluronan-
37	410	17.5	810	1 PLHU	plasma (EC 3.4.21
38	409.5	17.5	282	2 I84621	coagulation factor
39	406	17.4	417	1 S00845	hepsin (EC 3.4.21
40	404.5	17.3	275	2 A35863	trypsin (EC 3.4.2
41	404	17.3	1019	2 A38738	coagulation factor
42	403.5	17.3	275	2 B35863	trypsin (EC 3.4.2
43	403	17.2	583	2 A29154	trypsin (EC 3.4.2
44	402	17.2	246	2 B25528	complement factor
45	400.5	17.1	246	2 JQ1471	trypsin (EC 3.4.21
46	400	17.1	239	2 G42696	thrombin (EC 3.4.2
47	400	17.1	810	2 B30848	plasma (EC 3.4.21
48	399.5	17.1	231	1 TRFGTR	trypsin (EC 3.4.21
49	399	17.1	274	2 JC4171	trypsin (EC 3.4.2
50	399	17.1	1019	1 A56318	enteropeptidase (E
51	398	17.0	235	2 D42696	thrombin (EC 3.4.2
52	398	17.0	1047	2 A56117	masquerade precurs
53	397.5	17.0	236	2 C42696	thrombin (EC 3.4.2
54	396.5	17.0	270	2 S56160	mast cell tryptase
55	396.5	17.0	275	2 A32410	trypsin (EC 3.4.2
56	396.5	17.0	276	2 A38654	mast cell proteina
57	396	16.9	235	2 H42696	thrombin (EC 3.4.2
58	395	16.9	246	1 TRRT2	trypsin (EC 3.4.21
59	392.5	16.8	251	2 PC1235	29K serine protein
60	392	16.8	264	2 I38136	chymotrypsin-like

# ALIGNMENTS

## RESULT 1

KXHU

protein C (activated) (EC 3.4.21.69) precursor - human  
N:Alternate names: autoprothrombin IIA; plasma protein C  
C:Species: Homo sapiens (man)  
C>Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999  
C:Accession: A22331; A25426; A21781; A23789; A00927  
R:Foster, D.C.; Yoshitake, S.; Davie, E.W.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985  
A:Title: The nucleotide sequence of the gene for human protein C.  
A:Reference number: A22331; MUID:85270390; PMID:2991887  
A:Accession: A22331  
A:Molecule type: DNA  
A:Residues: 1-461 <FOS1>  
A:Cross-references: GB:M11228; NID:g190333; PIDN:AAA60166.1; PID:g190334  
R:Plutsky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.  
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986  
A:Title: Evolution and organization of the human protein C gene.  
A:Reference number: A25426; MUID:86120978; PMID:3511471  
A:Accession: A25426  
A:Molecule type: DNA  
A:Residues: 1-445, 'L', 446-461 <PLU>  
A:Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332



QY 1201 GACTGGATCATGGGCACATCAGACAAAGAGAGCCGCCAGAG 1245  
 |||||  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
 |||||  
 RESULT 2  
 KXBO  
 protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)  
 N;Alternate names: autoprothrombin IIA; plasma protein C  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 30-Nov-1980 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999  
 C;Accession: A26250; A18385; A18386; A00928  
 R;Long, G.L.; Balagaje, R.M.; MacGillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984  
 Title: Cloning and sequence of liver cDNA coding for bovine protein C.  
 Reference number: A26250; MUID:85014826; PMID:5091100  
 Accession: A26250  
 Molecule type: mRNA  
 A;Residues: 1-456 <LON>  
 R;Fennlund, P.; Stenflo, J.  
 J. Biol. Chem. 257, 12170-12179, 1982  
 A;Title: Amino acid sequence of the light chain of bovine protein C.  
 A;Reference number: A18385; MUID:83007325; PMID:6896876  
 A;Accession: A18385  
 A;Molecule type: protein  
 A;Residues: 40-194 <FER>  
 A;Note: 82-Lys was also found  
 R;Brakenberg, T.; Fennlund, P.; Roepstorff, P.; Stenflo, J.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983  
 A;Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.  
 A;Reference number: A19316; MUID:83169769; PMID:6572939  
 A;Contents: annotation; revision to residue 110  
 R;Stenflo, J.; Fennlund, P.  
 J. Biol. Chem. 257, 12180-12190, 1982  
 A;Title: Amino acid sequence of the heavy chain of bovine protein C.  
 A;Reference number: A18386; MUID:83007326; PMID:6896877  
 A;Accession: A18386  
 A;Molecule type: protein  
 A;Residues: 197-454, 'pv' <STX>  
 R;Esmon, N.L.; DeBault, L.E.; Esmon, C.T.  
 J. Biol. Chem. 258, 5548-5553, 1983  
 A;Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless F  
 A;Reference number: A37541; MUID:83213513; PMID:6304092  
 A;Contents: annotation; activation; calcium binding  
 R;Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.  
 J. Biol. Chem. 258, 5554-5560, 1983  
 A;Title: Structural changes required for activation of protein C are induced by Ca2+ bin  
 Reference number: A37542; MUID:83213514; PMID:6406503  
 A;Contents: annotation; activation; calcium binding  
 C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re  
 s.  
 C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c  
 bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this reacti  
 C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with stro  
 cognition of the thrombin-thrombomodulin complex.  
 C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin  
 C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
 C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding  
 F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>  
 F;24-83/Domain: Gla domain homology <GLA>  
 F;30-39/Domain: propeptide #status predicted <PRO>  
 F;40-194/Product: protein C light chain #status experimental <LCH>  
 F;98-128/Domain: EGF homology <EGF>  
 F;137-172/Domain: EGF homology <EG2>  
 F;197-456/Product: protein C heavy chain #status experimental <HCH>  
 F;197-210/Domain: activation peptide #status experimental <APT>  
 F;211-440/Domain: trypsin homology <TRY>  
 F;45-46, 53, 55, 58, 59, 62, 64, 65, 68, 74/Modified site: gamma-carboxyglutamic acid (Glu) #stat  
 F;110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
 F;119-128, 137-148, 144-157, 159-172, 180-318, 323-253, 368-382, 393-421/Disulfide bonds: #stat  
 F;136, 289, 350/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;252, 298, 397/Active site: His, Ser #status predicted  
 F;366/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:  
 Pred. No.: 1.08e-95 Length: 456  
 Score: 1660.00 Matches: 298  
 Percent Similarity: 80.96% Conservative: 38  
 Best Local Similarity: 71.81% Mismatches: 73  
 Query Match: 71.03% Indels: 6  
 DB: 1 Gaps: 2

US-09-763-153-1 (1-1245) x KXBO (1-456)

QY 1 GCCACTCTCTCTGGAGAGCTCGGTCCACAGCAGCCTGGAGCGGGAGTGCATAGAGGAG 60  
 |||||  
 Db 40 AlaAsnSerPheLeuGluGluLeuArgProGlyAsnValGluArgGluCysSerGluGlu 59  
 |||||  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 ::|||  
 Db 60 ValCysGluPheGluGluAlaArgGluIlePheGlnAsnThrGluAspThrMetAlaPhe 79  
 |||||  
 QY 121 TGGTCCAAAGCAGTCGACGGTGACAGTGTGGTCTTGCCCTTGGAGACCCCGTGGCC 180  
 |||||  
 Db 80 TrpSerPheTyrSerAspGlyAspGlnCysGluAspArgProSerGlySerProCysAsp 99  
 |||||  
 QY 181 AGCCTGTGCTCGGGCAGCGCAGCTGCATCGACGCGCATCGGCAGCTTCAGCTCGCACTGC 240  
 |||||  
 Db 100 LeuProCysCysGlyArgGlyLysCysIleAspGlyLeuGlyPheArgCysAspCys 119  
 |||||  
 QY 241 CGCAGCGCTGGAGGGCGCTTCTGCGAGCGGAGGTGAGCTTCTCTCAATGTCGCTG 300  
 |||||  
 Db 120 AlaGluGlyTrpGluGlyArgPheCysLeuHisGluValArgPheSerAspCysSerAla 139  
 |||||  
 QY 301 GACACGCGCGCTGCACGCATTACTGCTAGAGAGGTGGCGTGGCGGCTGTAGCTGT 360  
 ::|||  
 Db 140 GluAsnGlyGlyCysAlaHisTyrCysMetGluGluGlyArgArgHisCysSerCys 159  
 |||||  
 QY 361 GCGCTGTGCTACAAGCTGGGGCAGCACCTCTCGTGCAGTGTCAACCCGCGAGTGAAGTCCCT 420  
 |||||  
 Db 160 AlaProGlyTyrArgLeuGluAspHisGlnLeuCysValSerLysValThrPhePro 179  
 |||||  
 QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAGCGGAGTCACTGACCTGAACAGCAGACAGAA 480  
 |||||  
 Db 180 CysGlyArgLeuGlyLysArgMetGluLysLysArgLysThrLeuLysArgAspThrAsn 199  
 |||||  
 QY 481 -----GACCAAGAGACCAAGTAGATCCGCGGCTCATTGATGGAGAGATCACCAGCGG 534  
 |||||  
 Db 200 GlnValAspGlnLysAspGlnLeuAspProArgIleValAspGlyGlnGluAlaGlyTrp 219  
 |||||  
 QY 535 GGAGACAGCCCTCGCAGGTGCTCTGCTGGACTCAAGAGAGAGAGTGGCGCTCGGGGCA 594  
 |||||  
 Db 220 GlyLysSerProTrpGlnAlaValLeuLeuAspSerLysLysLysLeuValCysGlyAla 239  
 |||||  
 QY 595 GTGCTATCCACCCCTCTCGGTGTGTCAGACGGGCCCTCATGTCATGATGAGTCAAGAG 654  
 |||||  
 Db 240 ValLeuIleHisValSerTrpValLeuThrValAlaHisCysLeuAspSerArgLysLys 259  
 |||||  
 QY 655 CTCCTTCAGCGCTGGAGAGTATGACCTGGCGGCTGGGAGAGAGTGGGAGCTGGACCTG 714  
 |||||  
 Db 260 LeuIleValArgLeuGlyLysArgMetArgTrpGluSerTrpGluValAspLeu 279  
 |||||  
 QY 715 GACATCAAGGAGGTCTTCGTCACCCCACTACAGAGAGAGCAGCACCACGACATGACATC 774  
 |||||  
 Db 280 AspIleLysGluValIleIleHisProAsnTyrThrLysSerThrSerAspAsnIle 299  
 |||||  
 QY 775 GCATGTGTCACCTGGGCCCGCCACCTCTCGCAGACCATAGTGCCCATCTGCTC 834  
 |||||  
 Db 300 AlaLeuLeuArgLeuAlaLysProAlaThrLeuSerGlnThrIleValProIleCysLeu 319  
 |||||  
 QY 835 CCGGACAGCGCGCTTCGAGCGGAGCTCAATCAGCGCGCCGAGGAGACCTCGTACG 894  
 |||||  
 Db 320 ProAspSerGlyLeuSerGluArgLysLeuThrGlnValGlnGlnThrValValThr 339  
 |||||  
 QY 895 GGCTGGGCTACACAGCAGCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 954  
 |||||







[illegible]

QY 982 CACAATGAGTGCAGCGAGGTCATGAGCAACATGGTGTCTGACAACATGCTGTGTGCGGGC 1041

QY 982 CACAATGAGTGCAGCGAGGTCATGAGCAACATGGTGTCTGACAACATGCTGTGTGCGGGC 1041

Db	385	ArgAsnThrCysArgLeuSerThrSerPheSerIleThrGlnAsnMetPheCysAlaGly	404
QY	1042	ATCCTCGGGAGCCGACAGGATGCTCGAGGGCGACAGTGGGGGGCCATCGTGCCTCC	1101
Db	405	TyrAspAlaLysGlnGluAspAlaCysGlnGlyAspSerGlyGlyProHisValThrArg	424
QY	1102	TTCCACGGCACCTGGTTCCTGGTGGGCTGGTGAAGCTGGGGTGAAGGCTGGGGCTT	1161
Db	425	PheLysAspThrTyrPheValThrGlyIleValSerTrpGlyGluGlyCysAlaArgLys	444
QY	1162	CACAACTTACGGCGTTTACACAAAGTCAGCGCTACCTCGACTGGATGCCACATC	1221
Db	445	GlyLysTyrGlyIleTyrThrLysValThrAlaPheLeuLysTrpIleAspArgSerMet	464
	1222	AGAGACAAAGGAGGCCCC	1239
	465	LysAlaArgValGlyPro	470

## RESULT 6

## EXBO

coagulation factor Xa (EC 3.4.21.6) precursor - bovine

NA:Alternate names: Stuart factor

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence-revision 17-Mar-1987 #text-change 16-Jul-1999

C:C:Accession: A32867; A14997; A12030; A34412; S39414; A00925

R:M:Campbell, R.M.; MacGillivray, T.A.

N:Nucleic Acids Res 12, 4481-4492, 1984

A:A>Title: Blood coagulation factor X mRNA encodes a single polypeptide chain containing a

A:A:Reference number: A32867; MUID:84247315; PMID:6330671

A:A:Accession: A22867

A:A:Molecule type: mRNA

A:A:Residues: 1-487 <F>N

R:R:Enfield, D.L.; Ericsson, L.H.; Fujikawa, K.; Walsh, K.A.; Neurath, H.; Titani, K.

B:Biochemistry 19, 659-667, 1980

A:A>Title: Amino acid sequence of the light chain of bovine factor X-1 (Stuart factor).

A:A:Reference number: A14997; MUID:80130563; PMID:6766735

A:A:Accession: A14997

A:A:Molecule type: protein

A:A:Residues: 41-102, 'N', 104-180 <ENF>

R:R:McMullen, B.A.; Fujikawa, K.; Kisiel, W.

B:Biochem. Biophys. Res. Commun. 115, 8-14, 1983

A:A>Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co

A:A:Reference number: A20274; MUID:83308813; PMID:6688526

A:A:Contents: annotation; revision to residue 103

R:R:Titani, K.; Fujikawa, K.; Enfield, D.L.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.

O:O: Natl. Acad. Sci. U.S.A. 72, 3082-3086, 1975

A:A>Title: Bovine factor X-1 (Stuart factor): amino-acid sequence of heavy chain.

A:A:Reference number: A12030; MUID:76053069; PMID:1059093

A:A:Accession: A12030

A:A:Molecule type: protein

A:A:Residues: 183-292, 294-295, 'GDE', 299-334, 336-348, 'AE', 351-354, 356-441, 'CKFG', 446-492 <T>

A:A:Note: carboxylate binding sites and disulfide bonds were determined

R:R:Persson, E.; Seltander, M.; Linse, S.; Drakenberg, T.; Oehlin, A.K.; Stenflo, J.

J:J: Biol. Chem. 264, 16897-16904, 1989

A:A>Title: Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal

A:A:Reference number: A34412; MUID:89380326; PMID:2789221

A:A:Accession: A34412

A:A:Molecule type: protein

A:A:Residues: 85-126 <PER>

A:A:Note: beta-hydroxyaspartic acid site

R:R:Inoue, K.; Morita, T.

J:J: Biochem. 218, 153-163, 1993

A:A>Title: Identification of O-linked oligosaccharide chains in the activation peptides of

A:A:Reference number: S39414; MUID:94062825; PMID:8243461

A:A:Accession: S39414

A:A:Molecule type: protein

A:A:Residues: 183-196; 199-209; 216-233 <INO>

A:A:Note: carboxylate binding sites

R:R:Titani, K.; Hermodson, M.A.; Fujikawa, K.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.; D

B:Biochemistry 11, 4899-4903, 1972

A:A>Title: Bovine factor X-1a (activated Stuart factor). Evidence of homology with mammali

A:A:Reference number: A12435; MUID:73053314; PMID:4264286

Db 97 Asn-----GlnGlyHisCysLysAspGlylleGlyAspTyrThrCysThrCys 112  
QY 241 CGCAGCGCTGGGAGCGGCTTCTGCCAG-----CCGAGGTCAGCTTCTCCTCAAT 291  
Db 113 AlaGluGlyPheGluGlyLysAsnCysGlnPheSerThrArgGluIle----- 128  
QY 292 TGCTCGCTGGACACAGCGGCTGCACCGATTACTGCTAGAGGAGTGGCTGGCGGCGC 351  
Db 129 CysSerLeuAspAsnGlyGlyCysAspGlnPheCysArgGluGluArgSerGluValArg 148  
QY 352 TGCTAGCTCGCTGGCTCAAGCTGGGAGCAGCCTCTGTCAGTGTACCCCGCAGTG 411  
Db 149 CysSerCysAlaHisGlyTyrValLeuGlyAspSerLysSerCysValSerThrGlu 168  
QY 412 AAGTTCCCTTGTGGAGG---CCCTGGAAGCGGATGGAGAGAGCAGCTACCTGAAA 468  
Db 169 ArgPheProCysGlyLysPheThrGlnGlyArgSerArgTyrAlaIleHisThrSer 188  
QY 469 CGAGACACAGAAGAC-----CAGAGACCAAGTAGATCGG----- 504  
Db 189 GluAspAlaLeuAspAlaSerGluLeuGluHisTyrAspProAlaAspLeuSerProThr 208  
QY 504----- 504  
Db 209 GluSerSerLeuAspLeuGlyLeuAsnArgThrGluProSerAlaGlyGluAspGly 228  
QY 505-----CGGCTCATGTATGGGAAGATGACCGCGGGGAGACAGCCCTGGCAG 552  
Db 229 SerGlnValValArgIleValGlyClyArgAspCysAlaGluGlyGlyCysProTyrPhe 248  
QY 553 GTGGTCTGTGGACTCAAGAGAGAGCTGCGCTGCGGGGAGCTGCTATCCACCCCTCC 612  
Db 249 AlaLeuLeuValAsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuAsnGluPhe 268  
QY 613 TGGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 672  
Db 269 TyrValLeuThrAlaAlaHisCysLysHisGlnAlaLysArgPheThrValArgValGly 288  
QY 673 GAGTATGACCTCGCGCGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGAGAGTCTTC 732  
Db 289 AspArgAsnThrGluGlnGluGluGlyAsnGluMetalHisGluValGluMetThrVal 308  
QY 733 GTCCACCCCACTACAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 792  
Db 309 LysHisSerArgPheValLysGluThrTyrAspPheAspIleAlaValLeuArgLeuLys 328  
QY 793 CAGCCCGCCACCTCTCCGACACCATAGTGCCTCTGCTCCCGGAGCAGCGGCTTGCA 852  
Db 329 ThrProIleArgPheArgAsnValAlaProAlaCysLeuProGluLysAspTyrAla 348  
QY 853 GAGCCGAGCTCATCAGCGCGGCGGAGGAGC---CTGCTGAGCGGCTGGGCTACAC 909  
Db 349 GluAlaThrLeu-----MetThrGlnLysThrGlyIleValSerGlyPheGly----- 364  
QY 910 AGCAGCCGAGAGAGAGGAGGCGCAAGAAACCGCACCTTCTGCTCCCACTTCATCAAGATT 969  
Db 365-----ArgThrHisGluLysGlyArgLeuSerSerThrLeuLysMetLeuGluVal 381  
QY 970 CCGGTGTCGCGCAATAGTACGAGGAGGATGATGACCAACATGCTGTCTGAGAACATG 1029  
Db 382 ProTyrValAspArgSerThrCysLysLeuSerSerPheThrIleThrProAsnMet 401  
QY 1030 CTGCTGCGGCGATCTCTCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1089  
Db 402 PheCysAlaGlyTyrAspThrGlnProGluAspAlaCysGlnGlyAspSerGlyGlyPro 421  
QY 1090 ATGGTGGCTCTCTCCAGCGCACCTGCTCTGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGG 1149  
Db 422 HisValThrArgPheLysAspThrTyrPheValThrGlyIleValSerTyrPheGlyGly 441  
QY 1150 TGTGGGCTCTCTCAACATCGGGGTTTACACCAAGTCAGCCGCTACCTCGATGGATC 1209  
Db 442 CysAlaArgLysGlyLysPheGlyValTyrThrLysValSerAsnPheLeuLysTrpIle 461

QY 1210-----CATGGGCACATCAGACACAG 1230  
Db 462 AspLysIleMetLysAlaArgAlaGlyAlaAlaGlySerArgGlyHis-----Ser 478  
QY 1231 GAAGCCCCC 1239  
Db 479 GluAlaPro 481  
RESULT 7  
EXCH  
coagulation factor Xa (EC 3.4.21.6) precursor - chicken  
N;Alternate names: virus-activating proteinase  
C;Species: Gallus gallus (chicken)  
C;Date: 12-Feb-1993 #sequence\_revision 07-Feb-1997 #text\_change 16-Jul-1999  
C;Accession: S15838; S20380; S20381  
R;Suzuki, H.; Harada, A.; Hayashi, Y.; Wada, K.; Asaka, J.; Gotoh, B.; Ogasawara, T.;  
FEBS Lett. 283, 281-285, 1991  
A;Title: Primary structure of the virus activating protease from chick embryo. Its id  
A;Reference number: S15838; MUID:91257322; PMID:2044767  
A;Accession: S15838  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-475 <SUZ>  
A;Cross-references: DDBJ:D00844; NID:g222869; PIDN:BAA00724.1; PID:g222870  
R;Gotoh, B.; Yamauchi, F.; Ogasawara, T.; Nagai, Y.  
FEBS Lett. 296, 274-278, 1992  
A;Title: Isolation of factor Xa from chick embryo as the amniotic endoprotease respon  
A;Reference number: S20380; MUID:92164779; PMID:1537403  
A;Accession: S20380  
A;Molecule type: protein  
A;Residues: 41-55 <G02>  
A;Accession: S20381  
A;Molecule type: protein  
A;Residues: 241-246,'X',248-251,'X',253-261 <G0T>  
C;Function:  
C;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the  
A;Pathway: blood coagulation  
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol  
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-40/Domain: propeptide #status predicted <PRO>  
F;25-84/Domain: Gla domain homology <GLA>  
F;1-185/Product: coagulation factor X light chain #status experimental <LCH>  
F;90-121/Domain: EGF homology <EG1>  
F;129-167/Domain: EGF homology <EG2>  
F;186-475/Product: coagulation factor X heavy chain #status predicted <HCH>  
F;186-240/Domain: activation peptide #status predicted <APT>  
F;241-475/Product: coagulation factor Xa heavy chain #status experimental <AHC>  
F;241-468/Domain: trypsin homology <TRY>  
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu)  
F;57-62,90-101,95-110,112-121,129-140,136-152,154-167,175-348,247-252,267-283,396-410  
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted  
F;196,207,228,285/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;282,328,425/Active site: His, Asp, Ser #status predicted

Alignment Scores: 2,47e-42 Length: 475  
Pred. No.: 801.50 Matches: 163  
Score: 54.77% Conservative: 84  
Percent Similarity: 36.14% Mismatches: 147  
Best Local Similarity: 34.30% Indels: 57  
Query Match: 1 Gaps: 8  
DB: 1

US-09-763-153-1 (1-1245) x EXCH (1-475)

QY 1 GCCAACTCTCTCTCGGAGGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
Db 41 AlaAsnSerPheLeuGluGluMetLysGlnGlyAsnGluArgGluCysAsnGluGlu 60  
QY 61 ATCTGTGACTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
Db 61 ArgCysSerLysGluGluAlaArgGluAlaPheGluAsnGluLysThrGluGluPhe 80



A:Molecule type: protein		
A:Residues: 41-179 <MCM>		
R:Inoue, K.; Morita, T.		
Eur. J. Biochem. 218, 153-163, 1993		
A:Title: Identification of O-linked oligosaccharide chains in the activation peptides of		
A:Reference number: S39414; MUID:94062825; PMID:8243461		
A:Accession: S39415		
A:Molecule type: protein		
A:Residues: 183-234 <INO>		
A:Note: glycosylation sites		
A:Note: identification and characterization of beta-hydroxyaspartic acid		
R:Jagadeeswaran, P.; Reddy, S.V.; Rao, K.J.; Hamsabhusan, K.; Lyman, G.		
Gene 84, 517-519, 1989		
A:Title: Cloning and characterization of the 5' end (exon 1) of the gene encoding human		
A:Reference number: I54051; MUID:90128299; PMID:2612918		
A:Accession: I54051		
A>Status: translation not shown; translated from GB/EMBL/DBJ		
A:Molecule type: DNA		
A:Residues: 1-23 <RES>		
A:Cross-references: GB:M33297; NID:gl03860; PIDN:AAA52636.1; PID:g553330		
Padmanabhan, K.; padmanabhan, K.P.; Tullinsky, A.; Park, C.H.; Bode, W.; Huber, R.; Bla		
Mol. Biol. 232, 947-966, 1993		
A:Title: Structure of human des(1-45) factor Xa at 2.2 angstroms resolution.		
A:Reference number: A49458; MUID:93360277; PMID:835279		
A:Contents: annotation; X-ray crystallography, 2.2 angstroms		
C:Comment: The two chains held together by one disulfide bond are formed from a single-c		
C:Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway) or		
C:Genetics:		
A:Gene: GDB:F10		
A:Cross-references: GDB:l19890; OMIM:227600		
A:Map position: 13q34-13q34		
A:Introns: 24/1; 77/3; 86/1; 124/1; 150/3; 249/3; 289/1		
A:Note: deficiency of this factor causes Stuart disease		
C:Function:		
A:Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr		
A:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology		
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglut		
F:1-23/Domain: signal sequence #status predicted <SIG>		
F:24-40/Domain: propeptide #status predicted <PRO>		
F:25-84/Domain: Gla domain homology <GLA>		
F:41-179/Product: coagulation factor X light chain #status experimental <LCH>		
F:90-121/Domain: EGF homology <EGF>		
F:129-164/Domain: EGF homology <EG2>		
F:183-488/Product: coagulation factor X heavy chain #status experimental <HCH>		
F:183-234/Domain: activation peptide #status experimental <APT>		
F:235-488/Product: coagulation factor Xa heavy chain #status experimental <ACT>		
F:235-462/Domain: trypsin homology <TRY>		
F:46,47,54,56,59,60,65,66,69,72,79/Modified site: gamma-carboxyglutamic acid (Glu) #stat		
F:57-62/Disulfide bonds: #status predicted		
F:90-101,95-110,112-121,129-140,136-149,151-164,172-342,241-246,261-277,390-404,415-443/		
F:103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental		
F:199,211/Binding site: carboxylate (Thr) (covalent) #status experimental		
F:21,231/Binding site: carboxylate (Asn) (covalent) #status experimental		
F:34-235/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat		
F:276,322,419/Active site: His, Asp, Ser #status experimental		
Alignment Scores:		
Pred. No.: 2,47e-42	Length: 488	
Score: 801.50	Matches: 159	
Percent Similarity: 54.55%	Conservative: 87	
Best Local Similarity: 35.25%	Mismatches: 150	
Query Match: 34.30%	Indels: 55	
DB: 1	Gaps: 8	
US-09-763-153-1 (1-1245) x EXHU (1-488)		
QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTCACAGCAGCGCTGGAGCGGAGTGTCATAGAGGAG 60		
Db 41 AlaAsnSerPheLeuGluMetLysLysGlyHisLeuGluArgGluCysMetGluGlu 60		
QY 61 ATCTGTGACATCGAGGAGGCGCAAGGAAATTTCCAAATGTGATGACACACTGGCCTTC 120		





F;91-127/Domain: EGF homology <EG2>  
F;153-407/Product: coagulation factor vIIa heavy chain #status experimental <MA2>  
F;153-387/Domain: trypsin homology <TRY>  
F;6,7,14,16,19,20,25,26,29,34,35/Modified site: gamma-carboxyglutamic acid (Glu) #sta  
F;17-22,50-61,55-70,72-81,91-102,98-112,114-127,135-262,159-164,178-194,310-329,340-53  
F;152/Binding site: carboxylate (Ser) (covalent) #status experimental  
F;63/Modified site: erythro-beta-hydroxyaspartic acid (Asp) (partial) #status experim  
F;145,203/Binding site: carboxylate (Asn) (covalent) #status experimental  
F;132-153/Cleavage site: Arg-ile (coagulation factor XIa) #status experimental  
F;193,242,344/Active site: His, Asp, Ser #status predicted  
F;290-291/Cleavage site: Arg-Gly (coagulation factor Xa) #status experimental

Alignment Scores:  
Pred. No.: 5,71e-41 Length: 407  
Score: 779,50 Matches: 166  
Percent Similarity: 54,89% Conservative: 64  
Best Local Similarity: 39,62% Mismatches: 150  
Query Match: 33,35% Indels: 39  
DB: 1 Gaps: 11

US-09-763-153-1 (1-1245) x KFB07 (1-407)

Qy 1 GCACACTCTCTCCGGAGGAGCTCCGTACACAGCCTGGAGCGGAGTGCATAGAGGAG 60  
Db 1 AlaAsnGlyPheLeuGluGluLeuLeuProGlySerLeuGluArgGluCysArgGluGlu 20  
Qy 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATTTGGATGACACACTGGCGTTC 120  
Db 21 LeucySerPheGluGluAlaHisGluIlePheArgAsnGluGluArgThrArgGlnPhe 40  
Qy 121 TGGTCCAAAGCAGCTGACGGTGACAGTGTGTGCTTGCCCTGGAGCACCCGTGGCC 180  
Db 41 TrpValSerTyrAsnAspGlyAspGlnCys-----AlaSer 52  
Qy 181 AGCCTGTGCTCGGGCAGCGCACGTGTCATCGCGCATCGCAGCTTCAGCTCGCACTGC 240  
Db 53 SerProCysGlnAsnGlyGlySerCysGluAspGlnLeuArgSerTyrIleCysPheCys 72  
Qy 241 CGCAGCGCTGGAGGGCGCTTCTGCAGCGCGAG---GTGAGCTTCCTCAATTGCTCG 297  
Db 73 ProAspGlyPheGluGlyArgAsnCysGluThrAspLysGlnSerGlnLeuIleCysAla 92  
Qy 298 CTGGACAACGGCGCTGCACGCTACTTCCTCCTAGAGAGTGGGC---TGGCGCGCGTGT 354  
Db 93 AsnAspAsnGlyGlyCysGluGlnTyrCysGlyAlaAspProGlyAlaGlyArgPheCys 112  
Qy 355 AGCTGTGCGCTTGGCTACAGCTGGGGAGCAGCTTCGTGAGTGTCACTCCCGCGAGTGAAG 414  
Db 113 TrpCysHisGluGlyTyrAlaLeuGlnAlaAspGlyValSerCysAlaProThrValGlu 132  
Qy 415 TTCCTCTGGGAGGCCCTCGAAGCGGATGAGAAGAGCGCAGTCACTCAAAACGAGAC 474  
Db 133 TyrProCysGlyLys---IleProValLeuGluLysArgAsnGly----- 146  
Qy 475 ACAGAAGACCAAGAAGACAGTAGATCCGGCGCTCATTTGATGGAGAAGATACCAGCGCG 534  
Db 147 -----SerLysProGlnGlyArgIleValGlyHisValCysProLys 161  
Qy 535 GGAGACAGCCCTCGCAGGTGTCTGCTGGACTCAAGAGAAGAGCTGGCTCGCGGGCA 594  
Db 162 GlyGluCysProTrpGln---AlaMetLeuLysLeuAsnGlyAlaLeuLeuCysGlyGly 180  
Qy 595 GTGCTCATCCACCTCTCGGTGTGACAGGGCGCCACTGCATGGATGAGTCCAAG--- 651  
Db 181 ThrLeuValGlyProAlaIleValValSerAlaAlaHisCysPheGluArgLeuArgSer 200  
Qy 652 -----AAGCTCTTGTACAGGTGTGGAGAGTAGTATGCTGCGCGCTGGGAGAAGTGGAG 705  
Db 201 ArgGlyAsnLeuThrAlaValLeuGlyGluHisAspLeuSerArgValGluGlyProGlu 220  
Qy 706 CTGGACCTGGACATCAAGAGGTCTTCGTCCACCCCACTCAGCAAGACGACCCAGC 765  
Db 221 GlnGluArgArgValAlaGlnIleLeuValProLysGlnTyrValProGlyClnThrAsp 240

```

QY 766 AATGACATCGCACTGCTGCACCTGGCCAGCCGCGCCACCTCTCGCAGACCATAGTGGCC 825
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 HisAspValAlaLeuLeuGlnLeuAlaGlnProValAlaLeuGlyAspHisValAlaPro 260
QY 826 ATCTGCTCCCGGACAGCGGCTTCGAGAGCGGCTCAATCAGCGCGCCAGGAGACC 885
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 LeuCysLeuProAspPheAlaAspGlnThrLeuAlaPheVal---ArgPheSer 279
QY 886 CTCGTGACGGCTGGGCTACACAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 945
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 AlaValSerGlyTrpGlyGlnLeuLeuGluArgGlyValThrAlaArgLys----- 296
QY 946 TTCGTCTCTCACTTCATCAAGATTCCCGTGGTCCCGCACAATGAGTCAGCGAGTGCATG 1005
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 -----LeuMetValValLeuValProArgLeuLeuThrGlnAspCysLeuGlnSer 314
1006 AGCAAC-----ATGGTGTCTGAGAACATGTGTGTGGCGGCATCTCGGG 1050
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 ArgGlnArgProGlyGlyProValValThrAspAsnMetPheCysAlaGlyTyrSerAsp 334
QY 1051 GACCGGAGGATGCTCGAGGGGACAGTGGGGGCGCCATGGTGGCTCTTCCAGCGC 1110
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
335 GlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrArgPheArgGly 354
QY 1111 ACCTGGTTCCTGGTGGGCTGGTGGGCTGAGGCTGGGCTGGGCTCTTCAACAATC 1170
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 ThrTrpPheLeuThrGlyValValSerTrpGlyGluGlyCysAlaAlaGlyHisPhe 374
QY 1171 GGCGTTTACACCAAGTCACCGCTACCTCGACTGGATCCAT-----GGGCAC 1218
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
375 GlyIleThrArgValSerArgTyrThrAlaTrpLeuArgGlnLeuMetGlyHis 393

RESULT 11
I46932
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 12-Feb-1999
C:Accession: I46932
R:Brothers, A.B.; Clarke, B.J.; Sheffield, W.P.; Blajchman, M.A.
Thromb. Res. 69, 231-238, 1993
A:Title: Complete nucleotide sequence of the cDNA encoding rabbit coagulation factor VII
A:Reference number: I46932; MUID:93190306; PMID:6383365
A:Accession: I46932
A>Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
Residues: 1-443 <BRO>
Cross-references: GB:S56300; NID:Q266294; PID:Q266295
Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
F:24-83/Domain: Gla domain homology <GLA>
F:89-120/Domain: EGF homology <EGF>
F:130-166/Domain: EGF homology <EG2>
F:192-425/Domain: trypsin homology <TRY>

Alignment Scores:
Pred. No.: 2, 4e-40 Length: 443
Score: 769.50 Matches: 168
Percent Similarity: 55.56% Conservative: 67
Best Local Similarity: 39.72% Mismatches: 145
Query Match: 32.93% Indels: 43
DB: 2 Gaps: 12

US-09-763-153-1 (1-1245) x I46932 (1-443)
QY 1 GCCAATCTCTCTGGAGGAGTCCGTCACAGCAGCCTGGAGGGAGTGCATAGAGAG 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 AlaAsnSerPheLeuGluLeuLeuArgProGlySerLeuGluArgGluCysLysGluGlu 59
QY 61 ATCTGTGACTTCGAGAGCCCAAGGAAATTTCCAAATGTGATGACACACTGGCCCTTC 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 LeuCysSerPheLeuGluAlaArgGluValPheGlnSerThrGluArgThrLysGlnPhe 79
QY 121 TGGTCCCAAGCAGTGGAGGTTGACCAAGTGCCTGGTCTTGGCTTGGACACCCGTCGCC 180

```

```

Db 80 TrpLeuThrTyrAsnAspGlyAspGlnCysAlaSerAsnPro----- 93
QY 181 AGCCTGTGCTGCGGCGACGCGACGTGCATCGACGGATCGGCAGCTTACGCTGCGACTGC 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 -----CysGlnAsnGlyGlySerCysGluAspGlnIleGlnSerTyrIleCysPheCys 111
QY 241 CGCAGCGGTGGGAGGCGGCTTCTGCCAGCGCAGGTGAGC---TTCCTCAATTCTGCTCG 297
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
112 LeuAlaAspPheGluGlyArgAsnCysGluLysAsnLysAsnAspGlnLeuIleCysMet 131
QY 298 CTGCAACAACGGCGCTGCACGCAATTACTGCCTAGACAGAGTGGGTGGCGGCGC---TGT 354
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 TyrGluAsnGlyGlyCysGluGlnTyrCysSerAspHisValGlySerGlnArgSerCys 151
QY 355 AGCTGTGCGCTGGCTACAGTGTGGGGGACGACCTCTCAGTGTCCACCCGACGTGAAG 414
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 ArgCysHisGluGlyTyrThrLeuLeuProAsnGlyValSerCysThrProThrValAsp 171
QY 415 TTCCTTGTGGAGGCGCTTGGAGCGGTGAGAGAGAGCGGAGTCACTGAAACGAGAC 474
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 TyrProCysGlyLysVal---ProAlaLeuGluLysArgGlyAla----- 185
QY 475 ACAGAAGACCAAGAAGACCAAGTAGATCCGCGCTCATTGATGGGAAGATGACAGGCGG 534
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 -----SerAsnProGlnGlyArgIleValGlyGlyLysValCysProLys 200
QY 535 GGAGACAGCCCTGGCAGGTGCTGCTGAGTCAAGAGAAAGTCCGCTCGCGGGGCA 594
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 GlyGluCysProTrpGlnAlaAlaLeuMetAsnGly---SerThrLeuLeuCysGlyGly 219
QY 595 GTGCTCATCCACCCCTCTCGGTGTGACAGCGGCGCCACTGCATGGATGAG--- 645
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 SerLeuLeuAspThrHisTrpValValSerAlaAlaHisCysPheAspLysLeuSerSer 239
QY 646 TCCAAGAAGCTCTTGTGAGGTATGACCTGCGCGCTGGGAGAGTGGGAG 705
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 LeuArgAsnLeuThrIleValLeuGlyGluHisAspLeuSerGluHisGlyGlyAspGlu 259
QY 706 CTGGACCTGGACATCAAGAGAGTCTTCCTCCACCCCACTACAGCAAGAGCACCACG 765
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
260 GlnValArgHisValAlaGlnLeuIleMetProAspLysTyrValProGlyLysThrAsp 279
QY 766 AATGACATCGCACTGCTGCACCTGGCCAGCGCGCCGCCACCTCTCGCAGACCATAGTGGCC 825
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 HisAspIleAlaLeuLeuArgLeuGlnProAlaAlaLeuThrAsnAsnValValPro 299
QY 826 ATCTGCTCTCCGCGACAGCGGCTTGCAGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACC 885
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 LeuCysLeuProGluArgAsnPheSerGluSerThrLeuAlaThrIle---ArgPheSer 318
QY 886 CTCGTGACGGCTGGGCGC-----TACCACAGACCGCGAGAGAGAGGAGCCCAAGAGA 936
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 ArgValSerGlyTrpGlyGlnLeuLeuTyrArgGlyAlaLeuAlaArgGlu----- 335
QY 937 AACCGCACCTTCTGCTCAACTTCATCAAGATTCCCGTGGTCCCGCACATGAGTGCAGC 996
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 -----LeuMetAlaIleAspValProArgLeuMetThrGlnAspCysVal 350
QY 997 CAGGTCTATG-----AGCAACATGTGTCTGAGAACATGCTGTGTGCGGCGC 1041
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 GluGlnSerGluHisAsnProGlySerProGluValThrGlyAsnMetPheCysAlaGly 370
QY 1042 ATCTCGGGGACCGCGAGGATCGCGAGGCGCAGAGTGGGGGCGCCATGGTCCGCTCC 1101
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
371 TyrLeuAspGlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrSer 390
QY 1102 TTCACGCGCACCTGTTCTGCTGGTGGCTGTGAGCTGGGTGAGGCTGTGGGCTCCTT 1161
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
391 TyrHisGlyThr---TyrLeuThrGlyValValSerTrpGlyGluGlyCysAlaArgVal 409
QY 1162 CACAACCTACGGCGTTTACACCAAGTGCAGCGCTTACCTCGACTGGATTCATGGGCACATC 1221
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 410 GlyHisValGlyValTyrThrArgValSerArgAspThrGluTrpLeuSerArgLeuMet 429

QY 1222 AGAGACAAG 1230

Db 430 ArgSerLys 432

# RESULT 12

A30351

coagulation factor IXa (EC 3.4.21.22) precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A30351; I46201

R:Evans, J.P.; Watzke, H.H.; Ware, J.L.; Stafford, D.W.; High, K.A.

Blood 74, 207-212, 1989

A:Title: Molecular cloning of a cDNA encoding canine factor IX.

A:Reference number: A30351; MUID:89323338; PMID:2752110

A:Accession: A30351

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-452 <EVA>

Cross-references: GB:M21757; NID:g972719; PIDN:AAA75006.1; PID:g163948

Xelrod, J.H.; Read, M.S.; Brinkhous, K.M.; Verma, I.M.

Jc. Natl. Acad. Sci. U.S.A. 87, 5173-5177, 1990

Title: Phenotypic correction of factor IX deficiency in skin fibroblasts of hemophilic

A:Reference number: I46201; MUID:90311364; PMID:2367529

A:Accession: I46201

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-452 <AXE>

Cross-references: GB:M33826; NID:g163949; PIDN:AAA30844.1; PID:g163950

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-40/Domain: propeptide #status predicted <PRO>

F:24-84/Domain: Gla domain homology <GLA>

F:41-452/Product: coagulation factor IX #status predicted <MAT>

F:90-121/Domain: EGF homology <EG1>

F:127-163/Domain: EGF homology <EG2>

F:218-445/Domain: trypsin homology <TRY>

F:46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s

F:57-62,90-101,95-110,112-121,127-138,134-148,150-163,171-326,243-259,373-387,398-426/D

F:258,306,402/Active site: His, Asp, Ser #status predicted

## Alignment Scores:

Pred. No.:	6,09e-40	Length:	452
Score:	763.00	Matches:	158
Percent Similarity:	53.97%	Conservative:	73
Best Local Similarity:	36.92%	Mismatches:	141
Query Match:	32.65%	Indels:	56
DB:	1	Gaps:	12

us-09-763-153-1 (1-1245) x A30351 (1-452)

13 CTGAGAGAGCTCGTCACAGCAGCCCTGGAGCGGAGTGCATAGAGAGATCTGTGACTTC 72

45 LeuGluGluPheValArgGlyAsnLeuGluArgGlyCysIleGluGlyCysSerPhe 64

QY 73 GAGGAGCCCAAGAAATTTCCAAAATGTGGATGACACTGGCCCTTCCTGGTCCCAAGCAC 132

65 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheTrpLysGlnTyr 84

QY 133 CTCACGGGTGACAGTGTGTGTCCTGGCCCTGGAGCACCCTGGCCAGCCCTGTGCTGC 192

85 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 96

QY 193 GGGCAGCGCAGTCATCGACGCATCGGAGCTTCAGCTGGACGTCCGACGCGCGCTGG 252

97 AsnAspGlyValCysLysAspAlaAsnSerTyrGluCysTrpCysArgAlaGlyPhe 116

QY 253 GAGGGCGCTTCGCCAGCGCGAGGTGAGCTTCCTCAATTCCTGCTGCAGCAACGCGCGC 312

117 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 133

QY 313 TGCACGCACTACTGCTCCTAGAGAGAGTGGCTGGCGCGCGC-----TGTAGCTGTGGC 363

Db 134 CysLysGlnPheCys-----LysLeuGlyProAspAsnLysValValCysSerCysThr 151

QY 364 CTTGGCTACAAGCTGGGGGACGACCTCCTGCAGCTGCACCCCGCAGTGAAGTTCCTCTGT 423

Db 152 ThrGlyTyrGlnLeuAlaGluAspGlnArgSerCysGluProAlaValProPheProCys 171

QY 424 GGGAGG-----CCCTGGAGCGGATGGAGAGAGCGAGTCACCTGAAACGAGAC 474

Db 172 GlyArgValSerValProHisIleSerMetThrArgThrArgAlaGluThrLeuPheSer 191

QY 475 ACAGAGACCAAGACAGACCAAGTAGAT----- 501

Db 192 AsnMetAspTyrGluAsnSerThrGluValGluLysIleLeuAspAsnValThrGlnPro 211

QY 502 -----CCGCGGCTCATTGATGGAGAGATCACCAGCGGGGAGACAGCCCTGG 549

Db 212 LeuAsnAspPheThrArgValValGlyGlyLysAspAlaLysProGlyGlnPheProTrp 231

QY 550 CAGGTGTCTGCTGCTGACTCAAGAAGAGCTGGCTGGCGGCGAGTGTCTATCCACCCC 609

Db 232 Gln--ValLeuLeuAsnGlyLysValAspAlaPheCysGlySerIleIleAsnGlu 250

QY 610 TCTGTGTGTGACAGCGGCCCACTGCATGGATGAGTCCCAAGAGCTCCTGTGTCAGGCT 669

Db 251 LysTrpValValThrAlaAlaHisCysIleGluProAspValLysIleThrIleValAla 270

QY 670 GGAGATGATCACCTGGCGGCTGGGAGAGTGGACCTGGACATCAAGAGAGTTC 729

Db 271 GlyGluHisAsnThrGluLysArgGluHisThrGluGlnLysArgAsnValIleArgThr 290

QY 730 TTCCTCCACCCCACTACACAGACAGCACACC-----CACAAATGACATCGCAGCTGCTG 783

Db 291 IleLeuHisHisSerTyrAsnAlaThrIleAsnLysTyrAsnHisAspIleAlaLeuLeu 310

QY 784 CACCTGCCCGCAGCCCGCCACTCTCCAGACCATAGTCCCTCTCCCTCCCGGACAGC 843

Db 311 GluLeuAspGluProLeuThrLeuAsnSerTyrValThrProIleCys----- 326

QY 844 GGCCTTCAGAGCGCGAGCTCAATCAG-----GCCGCCAGGAGAGACC 885

Db 327 ---IleAlaAspArgGluTyrSerAsnIlePheLeuLysPheGlySerGly----- 342

QY 886 CTCGTGAGCGGCTGGCTACCCAGCAGCGGAGAGAGGAGCCCAAGANACCGCACC 945

Db 343 TyrValSerGlyTrpGlyArgValPheAsnLysGlyArgSerAlaSer----- 358

QY 946 TTCCTCTCAACTTCATCAAGATTCCCGTGGTCCCGCACCAATGAGTGCAGCGAGGTCA 1005

Db 359 ---IleLeuGlnTyrLeuLysValProLeuValAspArgAlaThrCysLeuArgSerThr 377

QY 1006 AGCAACATGGTCTGAGAACATCTGTGTGGGGATCTTCGGGGGACCGCAGGATGCC 1065

Db 378 LysPheThrIleTyrAsnAsnMetPheCysAlaGlyPheHisGluGlyGlyLysAspSer 397

QY 1066 TCGAGGCGCAGTGGGGGCGCCATGGTCCCTCTCCACGCGCAGCTGGTTCCTGTGTG 1125

Db 398 CysGlnGlyAspSerGlyGlyProHisValThrGluValGluGlyIleSerPheLeuThr 417

QY 1126 GGCCTGTGTGAGTGGGTGAGGGCTGTGGGCTCCTTCCAACTACGGCGTTTACACCAA 1185

Db 418 GlyIleIleSerTrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleThrLys 437

QY 1186 GTCAGCCCTACCTCGACTGGATC 1209

Db 438 ValSerArgTyrValAsnTrpIle 445

## RESULT 13

KFHU

coagulation factor IXa (EC 3.4.21.22) precursor [validated] - human

N:Alternate names: antihemophilic factor B; Christinas factor

C:Species: Homo sapiens (man)

C; Date: 17-Dec-1982 #sequence-revision 30-Jun-1987 #text-change 15-Sep-2000  
 A; Accession: A00922; A37570; A30511; A32989; A22673; A21337; A37546; A30623; A60486; A20  
 R; Yoshitake, S.; Schach, B.G.; Postner, D.C.; Davie, E.W.; Kurachi, K.  
 Biochemistry 24, 3736-3750, 1985  
 A; Title: Nucleotide sequence of the gene for human factor IX (antithemophilic factor B).  
 A; Reference number: A00922; MUID:86000558; PMID:2994716  
 A; Accession: A00922  
 A; Molecule type: DNA  
 A; Residues: 1-461 <YOS>  
 A; Cross-references: GB:K02402; NID:gl82612; PIDN:AA59620.1; PID:gl82613  
 R; Anson, D.S.; Choo, K.H.; Rees, D.J.G.; Giamelli, F.; Gould, K.; Huddleston, J.A.; Bro  
 EMBO J. 3, 1053-1060, 1984  
 A; Title: The gene structure of human anti-haemophilic factor IX.  
 A; Reference number: A37570; MUID:84236100; PMID:6329734  
 A; Accession: A37570  
 A; Molecule type: DNA  
 A; Residues: 1-461 <ANS>  
 A; Cross-references: GB:K02048  
 R; Reitsma, P.H.; Bertina, R.M.; Ploos van Amstel, J.K.; Riemsens, A.; Briet, E.  
 Blood 72, 1074-1076, 1988  
 A; Title: The putative factor IX gene promoter in hemophilia B Leyden.  
 A; Reference number: A30511; MUID:88327116; PMID:3416069  
 A; Accession: A30511  
 A; Molecule type: DNA  
 A; Residues: 8-24 <REI>  
 A; Cross-references: EMBL:X55008; NID:g311288; PIDN:CA838245.2; PID:g4469253  
 R; Koerber, D.D.; Bottema, C.D.K.; Buerstedde, J.M.; Sommer, S.S.  
 Am. J. Hum. Genet. 45, 448-457, 1989  
 A; Title: Functionally important regions of the factor IX gene have a low rate of polymor  
 A; Reference number: A32989; MUID:89371752; PMID:2773937  
 A; Accession: A32989  
 A; Status: not compared with conceptual translation  
 A; Molecule type: DNA  
 A; Residues: 30-92 <KOE>  
 R; McGraw, R.A.; Davis, L.M.; Noyes, C.M.; Lundblad, R.L.; Roberts, H.R.; Graham, J.B.; S  
 Proc. Natl. Acad. Sci. U.S.A. 82, 2847-2851, 1985  
 A; Title: Evidence for a prevalent dimorphism in the activation peptide of human coagulat  
 A; Reference number: A22673; MUID:85190593; PMID:3857619  
 A; Accession: A22673  
 A; Molecule type: mRNA  
 A; Residues: 1-193, 'T', 195-461 <MCG>  
 A; Cross-references: GB:M1309; NID:gl80552; PIDN:AAA52023.1; PID:gl80553  
 A; Note: The authors translated the codon ACA for residue 29 as Tyr  
 R; Jaye, M.; de la Salle, H.; Schamber, F.; Balland, A.; Kohli, V.; Findell, A.; Tolstosh  
 Nucleic Acids Res. 11, 2325-2335, 1983  
 A; Title: Isolation of a human anti-haemophilic factor IX cDNA clone using a unique 52-ba  
 A; Reference number: A21337; MUID:83220788; PMID:6687940  
 A; Accession: A21337  
 A; Molecule type: mRNA  
 A; Residues: 1-193, 'T', 195-461 <JAY>  
 A; Cross-references: GB:J00137; NID:gl82610; PIDN:AAA52763.1; PID:gl82611  
 R; Jagadeeswaran, P.; Lavellie, D.E.; Kaul, R.; Mohandas, T.; Warren, S.T.  
 Somat. Cell Mol. Genet. 10, 465-473, 1984  
 A; Title: Isolation and characterization of human factor IX cDNA: identification of Tag I  
 A; Reference number: A37546; MUID:84300526; PMID:6089357  
 A; Accession: A37546  
 A; Molecule type: mRNA  
 A; Residues: 38-193, 'T', 195-326 <JAG>  
 A; Cross-references: GB:M35672  
 R; Kurachi, K.; Davie, E.W.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6461-6464, 1982  
 A; Title: Isolation and characterization of a cDNA coding for human factor IX.  
 A; Reference number: A30623; MUID:83065193; PMID:6959130  
 A; Accession: A30623  
 A; Molecule type: mRNA  
 A; Residues: 1-12, 'S', 'G', 14-73, 'P', 75-82, 'K', 84-203, 'P', 205-216, 'G', 218-298, 'A', 299-356, 'A',  
 A; Cross-references: GB:J00136; NID:gl82608; PIDN:AAA98726.1; PID:gl82609  
 A; Experimental source: liver  
 R; Tharakan, J.; Strickland, D.; Burgess, W.; Drohan, W.N.; Clark, D.B.  
 Vox Sang. 58, 21-29, 1990  
 A; Title: Development of an immunoaffinity process for factor IX purification.  
 A; Reference number: A60486; MUID:90194857; PMID:2316207  
 A; Accession: A60486

A; Molecule type: protein  
 A; Residues: 47-52, 'XX', 55-60, 'X', 62, 'XX', 65 <THA>  
 R; McMullen, B.A.; Fujikawa, K.; Kisiel, W.  
 Biochem. Biophys. Res. Commun. 115, 8-14, 1983  
 A; Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood  
 A; Reference number: A20274; MUID:83308813; PMID:6688526  
 A; Accession: A20274  
 A; Molecule type: protein  
 A; Residues: 105-109, 'X', 111-115 <MCM>  
 R; Balland, A.; Faure, T.; Carvallo, D.; Cordier, P.; Ulrich, P.; Fournet, B.; de la S  
 Eur. J. Biochem. 172, 565-572, 1988  
 A; Title: Characterisation of two differently processed forms of human recombinant fac  
 A; Reference number: S02527; MUID:88166735; PMID:3280312  
 A; Accession: S02527  
 A; Molecule type: protein  
 A; Residues: 29-63 <BAL>  
 R; Jallat, S.; Perraud, F.; Dalemans, W.; Balland, A.; Dieterle, A.; Faure, T.; Meulle  
 EMBO J. 9, 3295-3301, 1990  
 A; Title: Characterization of recombinant human Factor IX expressed in transgenic mice  
 A; Reference number: S12058; MUID:91006024; PMID:2209546  
 A; Accession: S12058  
 A; Molecule type: mRNA; protein  
 A; Residues: 1-68 <JAL>  
 A; Note: processed forms expressed in recombinant system  
 R; Handford, P.A.; Baron, M.; Mayhew, M.; Willis, A.; Beesley, T.; Brownlee, G.G.; Cam  
 EMBO J. 9, 475-480, 1990  
 A; Title: The first EGF-like domain from human factor IX contains a high-affinity calc  
 A; Reference number: S12377; MUID:90151623; PMID:2406129  
 A; Accession: S12377  
 A; Molecule type: protein  
 A; Residues: 92-130 <HAN>  
 A; Note: NMR detection of calcium binding by domain expressed in recombinant system  
 R; de la Salle, C.; Charmantier, J.L.; Baas, M.J.; Schwartz, A.; Wiesel, M.L.; Grunbe  
 Thromb. Haemost. 70, 370-371, 1993  
 A; Title: A deletion located in the 3' non translated part of the factor IX gene respo  
 A; Reference number: I59612; MUID:94054330; PMID:8236150  
 A; Accession: I59612  
 A; Status: translated from GB/EMBL/DBDJ  
 A; Molecule type: DNA  
 A; Residues: 444-461 <RRS>  
 A; Cross-references: GB:S66752; NID:9439773; PIDN:AA828588.1; PID:9439774  
 R; Stoflet, E.S.; Koerber, D.D.; Sarkar, G.; Sommer, S.S.  
 Science 239, 491-494, 1988  
 A; Title: Genomic amplification with transcript sequencing.  
 A; Reference number: I59529; MUID:88127096; PMID:3340835  
 A; Accession: I59529  
 A; Status: translated from GB/EMBL/DBDJ  
 A; Molecule type: DNA  
 A; Residues: 290-359 <RE2>  
 A; Cross-references: GB:M19063; NID:gl82622; PIDN:AAA52456.1; PID:gl82623  
 R; Agarwala, K.L.; Kawabata, S.; Takao, T.; Murata, H.; Shimonishi, Y.; Nishimura, H.;  
 Biochemistry 33, 5167-5171, 1994  
 A; Title: Activation peptide of human factor IX has oligosaccharides O-glycosidically  
 A; Reference number: A54255; MUID:94227047; PMID:8172892  
 A; Accession: A54255  
 A; Molecule type: protein  
 A; Residues: 'D', 204, 'X', 206-211, 212, 'D', 214, 'X', 216-221, 'D' <AGA>  
 A; Note: the residues designated 'X' were determined to be threonine bound to carbohy  
 R; Di Scipio, R.G.; Kurachi, K.; Davie, E.W.  
 J. Clin. Invest. 61, 1528-1538, 1978  
 A; Title: Activation of human factor IX (Christmas factor).  
 A; Reference number: A18483; MUID:78194509; PMID:659613  
 R; McGraw, R.A.; Davis, L.M.; Noyes, C.M.; Graham, J.B.; Roberts, H.R.; Stafford, D.W.  
 Am. Soc. Hematol. Abstr. 64 (Suppl. 1), 262a, 1984  
 A; Reference number: A37569  
 A; Contents: annotation  
 A; Note: 194-Thr was also found  
 R; Morita, T.; Isaacs, B.S.; Esmon, C.T.; Johnson, A.E.  
 J. Biol. Chem. 259, 5698-5704, 1984  
 A; Title: Derivatives of blood coagulation factor IX contain a high affinity Ca2+-bind  
 A; Reference number: A37543; MUID:84185715; PMID:6425296

A:Contents: annotation; calcium binding  
 R:Morita, T.; Isaacs, B.S.; Esmon, C.T.; Johnson, A.E.  
 J. Biol. Chem. 260, 2583, 1985  
 A:Reference number: A37544  
 A:Contents: annotation; calcium binding, correction  
 R:Bentley, A.K.; Rees, D.J.G.; Rizza, C.; Brownlee, G.G.  
 Cell 45, 343-348, 1986  
 A:Title: Defective propeptide processing of blood clotting factor IX caused by mutation  
 A:Reference number: A37545; MUID:86189947; PMID:3009023  
 A:Contents: annotation; signal sequence cleavage site  
 R:Suchiro, K.; Kawabata, S.I.; Miyata, T.; Takeya, H.; Takamatsu, J.; Ogata, K.; Kamiya,  
 J. Biol. Chem. 264, 21257-21265, 1989  
 A:Title: Blood clotting factor IX B(M) Nagoya: substitution of arginine 180 by tryptophan  
 A:Reference number: A30622; MUID:90078229; PMID:2592373  
 A:Contents: annotation; sequence of mutant B(M) Nagoya  
 A:Note: carboxylation, glycosylation, and cleavage sites  
 R:Baron, M.; Norman, D.G.; Harvey, T.S.; Hanford, P.A.; Mayhew, M.; Tse, A.G.D.; Brownlee,  
 submitted to the Brookhaven Protein Data Bank, November 1991  
 A:Reference number: A51252; PDB:1LIXA  
 A:Contents: annotation; conformation by (1)H-NMR, residues 92-130

Note: recombinant form expressed in yeast  
 Comment: Factor IX is activated by factor Xla, which excises the activation peptide propeptide  
 Comment: The gamma-carboxyglutamic acid residues arise by posttranslational, vitamin K  
 Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with str

C:Genetics:  
 A:Gene: GDB:F9

A:Cross-references: GDB:119900; OMIM:306900

A:Map position: Xq27.1-Xq27.2

A:Introns: 30/1; 84/2; 93/1; 131/1; 174/1; 241/3; 280/1

C:Function:

A:Description: catalyzes the proteolytic activation of coagulation factor X in the presence of factor Xa

A:Pathway: blood coagulation intrinsic pathway

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamic acid

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-46/Domain: propeptide #status experimental <PP>

F:31-91/Domain: Gla domain homology <GLA>

F:47-191/Product: coagulation factor IXa light chain #status experimental <ALC>

F:97-128/Domain: EGF homology <EG1>

F:134-170/Domain: EGF homology <EG2>

F:192-226/Domain: activation peptide #status experimental <ACT>

F:227-461/Product: coagulation factor IXa heavy chain #status experimental <AHC>

F:53-54,61,63,66,67,72,73,76,79,82,86/Modified site: gamma-carboxyglutamic acid (Glu) #

F:64-69,97-108,102-117,119-128,134-145,141-155,157-170,178-335,252-268,382-396,407-435/

F:99/Binding site: carboxylate (Ser) (covalent) #status experimental

F:101/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

F:191-192/Cleavage site: Arg-Ala (coagulation factor IXa) #status experimental

F:203,213/Binding site: carboxylate (Asn) (covalent) #status experimental

F:205,215/Binding site: carboxylate (Thr) (covalent) #status experimental

F:226-227/Cleavage site: Arg-Val (coagulation factor IXa) #status experimental

Alignment Scores:

Seq. No.:	Length:	461
1	736.00	150
Percent Similarity:	52.36%	72
Best Local Similarity:	35.38%	156
Query Match:	31.49%	46
DB:	1	10

US-09-763-153-1 (1-1245) x KFHU (1-461)

QY 13 CTGAGGAGCTCCGTACACAGCCGAGGAGTGCATGAGGAGTCTGTGACTTC 72

Db 52 LeuGluGluPheValGlnGlnAsnLeuGluArgGluCysMetGluGluLysCysSerPhe 71

QY 73 GAGAGGCCAAGGAATTTCCAAATGTGGATGACACATCGCCCTCTGGTCCCAAGCAC 132

Db 72 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheThrLysGlnTyr 91

QY 133 GTCACGGTGACCACTGCTGGTCTTCCTGGAGCAGCCGCGCCAGCCCTGTCTGC 192

Db 92 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 103

QY 193 GGGCAGCGCAGGTGCATCGAGCGCATCGCGACTTCAGCTCGGACTGCCGACGCGCTGG 252

Db 104 AsnGlyCysCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 123

QY 253 GAGGCGCGCTTCTGCCAGCGCGAGGTGAGCTTCCTCAATTGCTGCTGCACACGCGCGC 312

Db 124 GluGlyAsnGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 140

QY 313 TGCACGCAATCTAGCTAGAGGAGGTGGGTGGCGG---CGCTGTAGCTGTGGCGCTGGC 369

Db 141 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGlu 160

QY 370 TACAAGCTGGGAGCAGCTCCCTGAGCTGTCACCCGCGAGTGAAGTCCCTTGTGGGAGG 429

Db 161 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 180

QY 430 CCTGGAGCGGATGAGAGGAGCGGAGTCACTCAACGCA----- 471

Db 181 ValSerValSerGlnThrSerLysLeuThrArgAlaGluAlaValPheProAspValAsp 200

QY 472 -----GACACAGAGCAAGCAAGCAAGCAAGTATAGT----- 501

Db 201 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer 220

QY 502 -----CGCGGCTCATTTGATGGAGAGTACCAGCGGGGAGACACCCCTGG 549

Db 221 PheAsnAspPheThrArgValValGlyGluAspAlaLysProGlyGlnPheProTrp 240

QY 550 CAGTGTCTCTGGTGGACTCAAGAGAACTGGCTGCGGCGGAGTGTCTATCCACCCC 609

Db 241 GlnValVal---LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu 259

QY 610 TCCTGGTGTGTACAGCGCGCCCTGATGATGAGTCAAGAGTCTCTTGTGAGGCTT 669

Db 260 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla 279

QY 670 GGAGATATGACCTGCGCGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGGAGTC 729

Db 280 GlyGluHisAsnIleGluGluThrGluHisThrGluLysArgAsnValIleArgIle 299

QY 730 TTCGTCCACCCCACTACAGCAAGAGCACCAACC-----GACATGATCGCATCTGT 783

Db 300 IleProHisHisAsnTyrAsnAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu 319

QY 784 CACCTGGCCCGCCCGCCCGCCCTCTCGCAGACCATAGTGCCTCTCCCTCCCGCAGC 843

Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 339

QY 844 GGCCTTGAGAGCGCGCTCAATCAGCGCGCCGAGGAGACCTCGTGCACGGCTGGGGC 903

Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 357

QY 904 -----TACCACAGCGCGAGAGAGGAGGAGCAAGAAACCGCACCTCTCTCTCAAC 957

Db 358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370

QY 958 TTCATCAAGATTCCTCGTGGTCCCGCCACAATGAGTCCACCGAGGTATGAGCAACATGTG 1017

Db 371 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 390

QY 1018 TCTGAGAACATCTGTGTGCGGCGCATCTCGGGGACCGGAGGAGTCTCGCAGGCGCAC 1077

Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyGlyArgAspSerCysGlnGlyAsp 410

QY 1078 AGTGGGGGCGCCATGGTGCCTCTCTCCACGGCCTGTTCTTCTGGTGGTGGTGGTGGT 1137

Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430

QY 1138 TGGGTGAGGGCTGGGCTCTTCCAACTACACGCGTTTACACCAAGTACGCGCTAC 1197

Db 431 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 450



QY 1198 CTCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454

## RESULT 14

JQ0419  
 coagulation factor IXa (EC 3.4.21.22) precursor - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 16-Jul-1999  
 C:Accession: JQ0419; 149667  
 R:Wu, S.W.; Stafford, D.W.; Ware, J.  
 Gene 86, 275-278, 1990

A:Title: Deduced amino acid sequence of mouse blood-coagulation factor IX.  
 A:Reference number: JQ0419; MUID:90215309; PMID:2323576

A:Accession: JQ0419

A:Molecule type: mRNA

A:Residues: 1-459 <WUS>

A:Cross-references: GB:M23109; NID:g193317; PIDN:AAA37629.1; PID:g387158

A:Experimental source: liver

R:Sarkar, G.; Koebler, D.D.; Sommer, S.S.

Genomics 6, 133-143, 1990

A:Title: Direct sequencing of the activation peptide and the catalytic domain of the factor IX.

A:Reference number: 146580; MUID:90152675; PMID:2303254

A:Accession: 149667

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 168-362, 'Q', 364-387, 'I', 389-451 <RES>

A:Cross-references: GB:M26236; NID:g193319; PIDN:AAA37630.1; PID:g193320

C:Comment: This protein plays a critical role in blood coagulation.

C:Superfamily: coagulation factor IX; EGF homology; Gla domain homology; trypsin homology

C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate

F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>

F:17-34/Domain: propeptide #status predicted <PRO>

F:19-79/Domain: Gla domain homology <GLA>

F:35-459/Product: coagulation factor IX #status predicted <MAT>

F:85-116/Domain: EGF homology <EGF>

F:122-158/Domain: EGF homology <EG1>

F:225-432/Domain: trypsin homology <TRY>

F:41-42,49,51,54,55,60,61,64,67,70,74/Modified site: gamma-carboxyglutamic acid (Glu) #

F:52-57,85-96,90-105,107-116,122-133,129-143,145-158,166-173,250-266,380-394,405-433/Dis

F:265,313,409/Active site: His, Asp, Ser #status predicted

## Alignment Scores:

Pred. No.: 1-22e-37 Length: 459  
 Score: 726.00 Matches: 155  
 Percent Similarity: 50.91% Conservative: 68  
 Local Similarity: 35.39% Mismatches: 151  
 Every Match: 31.07% Indels: 64  
 Gaps: 2

US-09-763-153-1 (1-1245) x JQ0419 (1-459)

QY 13 CTGAGGAGCTCCGTCACAGCAGCCTGGAGGAGTCATAGAGAGATCTGTGACTTC 72

Db 40 LeuGluGluPheValArgGlyAsnLeuGluGluCysIleGluGluArgCysSerPhe 59

QY 73 GAGAGCGCAAGGAATTTCCAAATGTGGATGACACTGGCTTCTGGTCCCAAGCAC 132

Db 60 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheTrpLysGlnTyr 79

QY 133 GTCCAGCGTGACCACTGCTTGGTCTTCCCTTGGAGCACCCGCTGGCGGACCTGTGCTGC 192

Db 80 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 91

QY 193 GGGCAGCGCAGTCATCGCAGCGCTCGGAGCTTCAGCTGCGACTGCGCCGACGCTGG 252

Db 92 AsnGlyGlyIleCysLysAspAspIleSerSerTyrGluCysTrpCysGlnValGlyPhe 111

QY 253 GAGGCGCGCTTGTCCCGCAGCGAGGTGAGCTTCAATGTCTGCTGACACACGCGCG 312

Db 112 GluGlyArgAsnGluLeuAspAlaThr-----CysAsnIleLysAsnGlyArg 128

QY 313 TGCACGCAATTACTGCCTAGAGAGGTGGCTGCGCG---CGCTGTAGCTGTGCGCGCTGCG 369

Db 129 CysLysGlnPheCysLysAsnSerProAspAsnLysValIleCysSerCysThrGluGly 148  
 QY 370 TACAAGCTGGGGAGAGACCTCTCTCAGTGTACCCCGGAGTGAAGTTCCTTGTGGGAGG 429  
 Db 149 TyrGlnLeuAlaGluAspGlnLysSerCysGluProThrValProPheProCysGlyArg 168  
 QY 430 CCC-----TGGAGCGGATG 444  
 Db 169 AlaSerIleSerTyrSerSerLysLysIleThrArgAlaGluThrValPheSerAsnMet 188  
 QY 445 GAGAGAGCGCATCTAC-----462  
 Db 189 AspTyrGluAsnSerThrGluAlaValPheIleGlnAspAspIleThrAspGlyAlaIle 208  
 QY 463 CTGAACAGAGACAGAGACCAAGAGACCAAGTAGAT---CCGCGCTCATGTGGG 519  
 Db 209 LeuAsnAsnValThrGluSerSerGluSerLeuAsnAspPheThrArgValValGlyGly 228  
 QY 520 AAGATGACCGGGGGGAGAGACGCCCTGGCAGGTGTCTGTGGACTCAAAAGAGAG 579  
 Db 229 GluAsnAlaLysProGlyGlnIleProTrpGlnValIle---LeuAsnGlyGluIleGlu 247  
 QY 580 CTGSCCTGGCGGCGAGTCTCATCCACCCCTCTGGTGTCTGACAGCGGCCCTCATG 639  
 Db 248 AlapheCysGlyGlyAlaIleAsnGluLysTrpIleValThrAlaAlaHisCysLeu 267  
 QY 640 GATGAGTCCAAGAGCTCTCTGTCAGGCTTGGAGAGTATGACCTGGCGGCTGGGAGAG 699  
 Db 268 LysProGlyAspLysIleGluValValAlaGlyGluThrAsnIleAspLysLysGluAsp 287  
 QY 700 TGGAGCTGACCTGGACATCAGGAGGTCTTCTGTCACCCCACTACAGACAGACACC 759  
 Db 288 ThrGluGlnArgArgAsnValIleArgThrIleProHisGlnTyrAsnAlaThrIle 307  
 QY 760 ACC-----GACAAATGACATCGCAGCTGTGACCTGGCGGCTGGCGGCTGGGAGAG 813  
 Db 308 AsnLysTyrSerHisAspIleAlaLeuLeuLysLysLysProLeuLeuLeuAsnSer 327  
 QY 814 ACCATAGTGCCTTCCTCCCGGACAGCGGCTTGCAGAGCGGAGCTCAATCAG---870  
 Db 328 TyrValThrProIleCys-----ValAlaAsnArgGluTyrThrAsnIle 342  
 QY 871 -----CGCGCCAGGAGACCTCTGTCAGGCTGGGGCTGGCGCTACACAGACG 915  
 Db 343 PheLeuLysPheGlySerGly-----TyrValSerGlyTrpGlyLysValPheAsn 359  
 QY 916 CGAGAGAAGGAGGCGCAAGAGAAACCGACCTCTGCTCTCAACTTCATCAAGATTCCCGTG 975  
 Db 360 LysGlyArgHisAlaSer-----IleLeuGlnTyrLeuArgValProLeu 374  
 QY 976 GTCCCGCCCAATGAGTGTGAGGAGGTCTATGAGCAACATGGTGTCTGAGAACATGTGTGT 1035  
 Db 375 ValAspArgAlaThrCysLeuArgSerThrThrPheThrThrTyrAsnAsnMetPheCys 394  
 QY 1036 GCGGGCTCTCGGGGACCGGAGGATGCTCGGAGGCGGAGTGGGGGGGCCCGGCTGTC 1095  
 Db 395 AlaGlyTyrArgGluGlyGlyLysAspSerCysGluGlyAspSerGlyGlyProHisVal 414  
 QY 1096 GCCTCTCTCCAGGCGACCTGTCTCTGTTGGGCTGTGGTGTGGTGGGCTGAGGCTGTGG 1155  
 Db 415 ThrGluValGluGlyThrSerPheLeuThrGlyIleIleSerTrpGlyGluGluCysAla 434  
 QY 1156 CTCCTTCAACATACGGCGTTTACACCAAGTCAACCGCTACCTGACCTGGATC 1209  
 Db 435 MetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyrTyrValAsnTrpIle 452

## RESULT 15

KFBO

coagulation factor IXa (EC 3.4.21.22) precursor - bovine

N:Alternate names: Christmas factor

C:Species: Bos primigenius taurus (cattle)

C>Date: 30-Nov-1980 #sequence\_revision 03-Aug-1984 #text\_change 16-Jul-1999

C;Accession: A14757; B20274; I45891; A00923  
R;Katayama, K.; Ericsson, L.H.; Enfield, D.L.; Walsh, K.A.; Neurath, H.; Davie, E.W.; T.  
Proc. Natl. Acad. Sci. U.S.A. 76, 4990-4994, 1979  
A;Title: Comparison of amino acid sequence of bovine coagulation factor IX (Christmas fa  
A;Reference number: A14757; MUID:80056619; PMID:291916  
A;Accession: A14757  
A;Molecule type: protein  
A;Residues: 1-63, 'T', 65-416 <KAT>  
R;McMullen, B.A.; Fujikawa, K.; Kistel, W.  
Biochem. Biophys. Res. Commun. 115, 8-14, 1983  
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co  
A;Reference number: A20274; MUID:83308813; PMID:668826  
A;Accession: B20274  
A;Molecule type: protein  
A;Residues: 59-63, 'X', 65-69 <MCM>  
R;Choo, K.H.; Gould, K.G.; Rees, D.J.G.; Brownlee, G.G.  
Nature 299, 178-180, 1982  
A;Title: Molecular cloning of the gene for human anti-haemophilic factor IX.  
A;Reference number: I45891; MUID:82272386; PMID:6287289  
A;Accession: I45891  
A;Title: Status: translated from GB/EMBL/DDBB  
A;Molecule type: mRNA  
A;Residues: 52-139 <CHO>  
Cross-references: GB:J00007; NID:g163053; PIDN:AAA30520.1; PID:g163054  
J;Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanaga,  
J. Biochem. 104, 867-868, 1988  
A;Title: A new trisaccharide sugar chain linked to a serine residue in bovine blood coag  
A;Reference number: A44556; MUID:89213999; PMID:3149637  
A;Contents: annotation  
A;Note: structure and location of a carbohydrate covalently bound to Ser  
C;Comment: Factor IX is activated by factor Xia, which excises the activation peptide ph  
C;Comment: The gamma-carboxyglutamic acid residues arise by posttranslational, vitamin K  
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with str  
C;Function:  
A;Description: catalyzes the proteolytic activation of coagulation factor X in the pres  
A;Pathway: blood coagulation intrinsic pathway  
C;Superfamily: coagulation factor X; BGF homology; Gla domain homology; trypsin homology  
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam  
F;1-146/Product: coagulation factor IXa light chain #status experimental <ALC>  
F;1-45/Domain: Gla domain homology (fragment) <GLA>  
F;51-82/Domain: EGF homology <EGL>  
F;88-124/Domain: EGF homology <EG2>  
F;147-181/Domain: activation peptide #status experimental <APT>  
F;182-416/Product: coagulation factor IXa heavy chain #status experimental <AHC>  
F;182-409/Domain: trypsin homology <TRY>  
F;7, 8, 15, 17, 20, 21, 26, 27, 30, 33, 36, 40/Modified site: gamma-carboxyglutamic acid (Glu) #sta  
F;18-23, 51-62, 56-71, 73-82, 88-99, 95-109, 111-124, 132-290, 207-223, 337-351, 362-390/Disulfide  
F;53/Binding site: carbohydrate (Ser) (covalent) #status experimental  
F;64/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
F;158, 168, 173, 261/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;222, 270, 366/Active site: His, Asp, Ser #status predicted

iment Scores:  
L. No.: 6, 27e-37 Length: 416  
714.50 Matches: 148  
Percent Similarity: 51.64% Conservative: 72  
Best Local Similarity: 34.74% Mismatches: 157  
Query Match: 30.57% Indels: 49  
DB: 1 Gaps: 10

US-09-763-153-1 (1-1245) x KFBO (1-416)

193 GGGACAGCGACCTGCTGATGAGCGGATCGGAGCTTACGCTGCGACTGCGCGAGCGGCTGG 252  
58 AsnGlyMetCysLysAspAspMetPheAsnSerValGluCysTrpCysGlnAlaGlyPhe 77  
253 GAGGCGCGCTTCTGCGACGCGGAGGTGAGCTTCTCAATTGCTCGCTGGACAAGCGGCGC 312  
78 GluGlyThrAsnCysGluLeuAspAlaThr-----CysSerIleLysAsnGlyArg 94  
313 TGCACGCTTACTGCTAGAGAGTGGCTGGCGG-----CGCTGTAGCTGTCGCGCTGC 369  
95 CysLysGlnPheCysLysArgAspThrAspAsnLysValValCysSerCysThrAspGly 114  
370 TACAAGCTGGGAGCAGACCTCTCCAGTGTCTACCCCGCAGTGAAGTTCCTTGTGGGAGG 429  
115 TyrArgLeuAlaGluAspGlnLysSerCysGluProAlaValProPheProCysGlyArg 134  
430 CCC-----TGGAGCGGATGGAGAGAGCGAGTCACTGAAACGAGAC 474  
135 ValSerValSerHisIleSerLysLysLeuThrArgAlaGluThrIlePheSerAsnThr 154  
475 ACAGAAGACCAAGAGACCAAGTAGAT----- 501  
155 AsnTyrGluAsnSerSerGluAlaGluIleIleTrpAspAsnValThrGlnSerAsnGln 174  
502 -----CCGCGCTCATTTGATGGGAAGATGACCGAGCGGGGAGACAGCCCC 546  
175 SerPheAspGluPheSerArgValValGlyGlyGluAspAlaGluArgGlyGlnPhePro 194  
547 TGCAGCTGGTCTGCTGTCAGTCAAGAGAGAGCTGGCTGGCGGCGGAGTGTCTATCCAC 606  
195 TrpGln---ValLeuLeuHisGlyGluIleAlaAlaPheCysGlyGlySerIleValAsn 213  
607 CCCTCTGGTGTGACAGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 666  
214 GluLysTrpValThrAlaAlaHisCysIleLysProGlyValLysIleThrValVal 233  
667 CTGGAGAGTATGCTGGCGCTGGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 726  
234 AlaGlyGluHisAsnThrGluLysProGluProThrGluGlnLysArgAsnValIleArg 253  
727 GTCTTCTGCTCCACCCCACTACAGCAAGAGCAGCCACC-----GACAAATGACATCGACTG 780  
254 AlaIleProTyrHisSerTyrAsnAlaSerIleAsnLysTyrSerHisAspIleAlaLeu 273  
781 CTGACCTGCCCGAGCGGCGGCTGCTGCGACGACATAGTGCCTATCTGCTCCCGGAC 840  
274 LeuGluLeuAspGluProLeuGluLeuAsnSerTyrValThrProIleCys----- 290  
841 AGCGGCTTTCAGAGCGGAGTCAATCAGCGCGGCGGAG-----ACCCCTCGTG 891  
291 -----IleAlaAspArgAspTyrThrAsnIlePheSerLysPheGlyTyrGlyTyrVal 308  
892 ACGGCTGGGGCTACACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951  
309 SerGlyTrpGlyLysValPheAsnArgGlyArgSerAlaSer-----Ile 323  
952 CTCAACTTCATCAAGATTCCTCGTGGTCCCGCACAATGAGTGCAGCGAGGTCATGAGCAAC 1011  
324 LeuGlnTyrLeuLysValProLeuValAspArgAlaThrCysLeuArgSerThrLysPhe 343  
1012 ATGTGTCTGAGAACATGCTGTGTCGGGCTCTCCGGGAGCGGAGGAGGAGGAGGAGGAGGAG 1071  
344 SerIleTyrSerHisMetPheCysAlaGlyTyrHisGluGlyLysAspSerCysGln 363  
1072 GGCACAGTGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1131  
364 GlyAspSerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIle 383  
1132 GTGAGCTGGGTGGGCTGTGGGCTCTCTCACAACATGCGGCTTTACACCAAGTCAAGC 1191  
384 IleSerTrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSer 403

QY 1192 CGTACCTCGACTGGATC 1209

|||||:|||||

Db 404 ArgTyrValAsnTrpIle 409

## RESULT 16

TBHU

Thrombin (EC 3.4.21.5) precursor [validated] - human  
 N:Alternate names: coagulation factor II  
 N:Contains: prothrombin

C:Species: Homo sapiens (man)

C:Date: 30-Nov-1980 #sequence\_revision 22-Jul-1994 #text\_change 08-Dec-2000  
 C:Accession: A29351; A00914; B00914; A37549; A37550; I51952

R:Degen, S.J.F.; Davie, E.W.

Biochemistry 26, 6165-6177, 1987

Title: Nucleotide sequence of the gene for human prothrombin.

Reference number: A29351; MUID:8807877; PMID:2825773

Accession: A29351

Molecule type: DNA

A:Residues: 1-622 &lt;DEG&gt;

A:Cross-references: GB:M17262; GB:M33691; NID:g558069; PIDN:AAC63054.1; PID:g339641

R:Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.

Biochemistry 22, 2087-2097, 1983

A:Title: Characterization of the complementary deoxyribonucleic acid and gene coding for

A:Reference number: A00914; MUID:83231469; PMID:6305407

A:Accession: A00914

Molecule type: mRNA

A:Residues: 8-163, 'N', 165-622 &lt;DE2&gt;

A:Cross-references: GB:V00595; GB:J00307; NID:g37128; PIDN:CAA23842.1; PID:g1335344

A:Accession: B00914

Molecule type: DNA

A:Residues: 188-311 &lt;DE3&gt;

R:Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.

Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977

A:Reference number: A37549; MUID:77193964; PMID:266717

A:Accession: A37549

Molecule type: protein

A:Residues: 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T', 184-193, 'MV', 196-308,

R:Butkowski, R.J.; Eliot, J.; Downing, M.R.; Mann, K.G.

J. Biol. Chem. 252, 4942-4957, 1977

A:Title: Primary structure of human prothrombin 2 and alpha-thrombin.

A:Reference number: A37550; MUID:77207112; PMID:873923

A:Accession: A37550

Molecule type: protein

A:Residues: 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413, 'N', 415-484, 'N', 486-

R:Rabiet, M.J.; Blaschill, A.; Furie, B.; Furie, B.C.

Biochemistry 26, 13210-13215, 1986

A:Reference number: A37551; MUID:87008532; PMID:3759958

Contents: annotation; activation cleavages

MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.

Ann. N. Y. Acad. Sci. 485, 73-79, 1986

A:Title: Recombinant genetic approaches to functional mapping of thrombin.

A:Reference number: I51952; MUID:87182874; PMID:34711151

A:Accession: I51952

Molecule type: mRNA

A:Residues: 1-2, 'RI', 5-100 &lt;RES&gt;

A:Cross-references: GB:M33031; NID:gl90723; PIDN:AAA60220.1; PID:gl90724

C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin

) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chain

ter 314-Arg, are released in natural blood clotting

C:Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.

C:Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxy

ent interaction with the negatively charged phospholipid membrane surface.

C:Comment: The prothrombin precursor is synthesized in the liver.

A:Gene: GDB:F2

A:Cross-references: GDB:I19894; OMIM:176930

A:Map position: 11p11-11q12

A:Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552

C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C:Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; dupli

F:1-24/Domain: signal sequence #status predicted &lt;SIG&gt;

F:25-43/Domain: propeptide #status predicted &lt;PRO&gt;

F:28-87/Domain: Gla domain homology &lt;GLA&gt;

F:44-622/Product: prothrombin #status experimental &lt;MAT&gt;

F:44-327/Domain: activation peptide #status experimental &lt;APT&gt;

F:108-186/Domain: kringle homology &lt;KRI&gt;

F:213-291/Domain: kringle homology &lt;KR2&gt;

F:328-363/Product: thrombin light chain #status experimental &lt;LCH&gt;

F:364-622/Product: thrombin heavy chain #status experimental &lt;HC&gt;

F:364-613/Domain: trypsin homology &lt;TRY&gt;

F:49,50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #sta

F:60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds: #sta

F:121,143/Binding site: carboxylate (Asn) (covalent) #status predicted

F:336-407/Disulfide bonds: #status predicted

F:391-407/Disulfide bonds: #status experimental

F:406,462/Active site: His, Asp #status predicted

F:416/Binding site: carboxylate (Asn) (covalent) #status experimental

F:568/Active site: Ser #status experimental

## Alignment Scores:

Pred. No.: 1.79e-27 Length: 622

Score: 562.50 Matches: 169

Percent Similarity: 39.76% Conservative: 62

Best Local Similarity: 29.09% Mismatches: 161

Query Match: 24.07% Indels: 189

DB: 1 Gaps: 24

US-09-763-153-1 (1-1245) x TBHU (1-622)

QY 1 GCCAACTCTCTCTCGAGAGCGTCGTCACACAGCGCTGGAGCGGAGTGATAGAGAG 60

Db 44 AlaAsnThrPheLeuGluValArgLysGlyAsnLeuGluArgGluCysValGluGlu 63

QY 61 ATCTGTGACTTCGAGAGCGCCAAAGAAATTTCCAAAATGTGGATGACACACTGGCCCTC 120

Db 64 ThrCysSerTyrGluAlaPheGluAlaLeuGluSerThrAlaThrAspValPhe 83

QY 121 TGGTCCACAGCAGCGTCGACAGTGCTGTGCTTGGCCCTGGACACCGCTGGGCC 180

Db 84 TrpAlaLysTyrThrAlaCysGluThrAlaArgThr---ProArgAspLysLeuAla 102

QY 181 AGCTGTGCTCGCGGCGACGCTGCATCGACGCGCATCGCGCAGC--- 225

Db 103 ---CysLeuGluGlyAsnCysAlaGluGlyLeuGlyThrAsnTyrArgGlyHis 119

QY 226 ---TTCAGCTGGCAC--- 237

Db 120 ValAsnIleThrArgSerGlyIleGluCysGlnLeuTrpArgSerArgTyrProHisLys 139

QY 238 ---TTCAGCTGGCAC--- 243

Db 140 ProGluIleAsnSerThrThrHisProGlyAlaAspLeuGlnGluAsnPheCysArgAsn 159

QY 244 ---AGCGGCTGGAGGCGCGCTCTGC--- 267

Db 160 ProAspSerSerThrThrGlyProTyrCysTyrThrThrAspProThrValArgArgGln 179

QY 268 ---CAGCGC--- 273

Db 180 GluCysSerIleProValCysGlyGlnAspGlnValThrValAlaMetThrProArgSer 199

QY 274 GAGGTGAGCTTCCTCAAT--- 312

Db 200 GluGlySerValAsnLeuSerProProLeuGluGlnCysValProAspArgGlyGln 219

QY 313 ---TGCACGCAT--- 321

Db 220 GlnTyrGlnGlyArgLeuAlaValThrThrHisGlyLeuProCysLeuAlaTrpAlaSer 239

QY 321 --- 321

Db 240 AlaGlnAlaLysAlaLeuSerLysHisGlnAspPheAsnSerAlaValGlnLeuValGlu 259

QY 322 ---TACTGCCTA--- 363



F:44-139/Domain: activation peptide 1 #status experimental <FR1>  
F:109-187/Domain: kringle homology <KR1>  
F:200-317/Domain: activation peptide 2 #status experimental <FR2>  
F:214-292/Domain: kringle homology <KR2>  
F:318-366/Product: thrombin light chain #status experimental <LCH>  
F:367-625/Product: thrombin heavy chain #status experimental <HCH>  
F:367-616/Domain: trypsin homology <TRY>  
F:501-51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental <G<  
F:61-66,91-104,109-187,130-170,158-182,214-292,233-275,263-287,339-485,394-410,539-553,581-600,611-625,636-650,661-675,686-700,711-726,737-752,763-777,788-803,814-829,840-855,866-881,892-907,918-933,944-959,970-985,996-1011,1022-1037,1048-1063,1074-1089,1090-1105,1116-1131,1142-1157,1168-1183,1194-1209,1220-1235,1246-1261,1272-1287,1298-1313,1324-1339,1350-1365,1376-1391,1402-1417,1428-1443,1454-1469,1480-1495,1506-1521,1532-1547,1558-1573,1584-1599,1600-1615,1626-1641,1652-1667,1678-1693,1704-1719,1730-1745,1756-1771,1782-1797,1808-1823,1834-1849,1860-1875,1886-1901,1912-1927,1938-1953,1964-1979,1990-2005,2016-2031,2042-2057,2068-2083,2094-2109,2120-2135,2146-2161,2172-2187,2198-2213,2224-2239,2250-2265,2276-2291,2302-2317,2328-2343,2354-2369,2380-2395,2406-2421,2432-2447,2458-2473,2484-2499,2510-2525,2536-2551,2562-2577,2588-2603,2614-2629,2640-2655,2666-2681,2692-2707,2718-2733,2744-2759,2770-2785,2796-2811,2822-2837,2848-2863,2874-2889,2895-2910,2921-2936,2947-2962,2973-2988,2994-3009,3020-3035,3046-3061,3072-3087,3098-3113,3124-3139,3150-3165,3176-3191,3202-3217,3228-3243,3254-3269,3280-3295,3306-3321,3332-3347,3358-3373,3384-3399,3410-3425,3436-3451,3462-3477,3488-3503,3514-3529,3540-3555,3566-3581,3592-3607,3618-3633,3644-3659,3670-3685,3696-3711,3722-3737,3748-3763,3774-3789,3795-3810,3821-3836,3847-3862,3873-3888,3894-3909,3920-3935,3946-3961,3972-3987,3998-4013,4024-4039,4050-4065,4076-4091,4102-4117,4123-4138,4149-4164,4175-4190,4201-4216,4227-4242,4253-4268,4279-4294,4305-4320,4331-4346,4357-4372,4383-4398,4409-4424,4435-4450,4461-4476,4487-4502,4513-4528,4539-4554,4565-4580,4591-4606,4617-4632,4643-4658,4669-4684,4695-4710,4721-4736,4747-4762,4773-4788,4794-4809,4820-4835,4846-4861,4872-4887,4893-4908,4919-4934,4945-4960,4971-4986,4992-5007,5018-5033,5044-5059,5070-5085,5096-5111,5122-5137,5148-5163,5174-5189,5195-5210,5221-5236,5247-5262,5273-5288,5294-5309,5320-5335,5346-5361,5372-5387,5398-5413,5424-5439,5450-5465,5476-5491,5502-5517,5523-5538,5549-5564,5575-5590,5601-5616,5627-5642,5653-5668,5674-5689,5695-5710,5721-5736,5747-5762,5773-5788,5794-5809,5820-5835,5846-5861,5872-5887,5893-5908,5919-5934,5945-5960,5971-5986,5992-6007,6018-6033,6044-6059,6070-6085,6091-6106,6117-6132,6143-6158,6164-6179,6190-6205,6216-6231,6242-6257,6268-6283,6294-6309,6315-6330,6341-6356,6362-6377,6383-6398,6404-6419,6430-6445,6456-6471,6482-6497,6503-6518,6524-6539,6545-6560,6571-6586,6592-6607,6613-6628,6634-6649,6655-6670,6681-6696,6702-6717,6723-6738,6744-6759,6765-6780,6791-6806,6812-6827,6833-6848,6854-6869,6875-6890,6901-6916,6922-6937,6943-6958,6964-6979,6985-7000,7011-7026,7032-7047,7053-7068,7074-7089,7100-7115,7121-7136,7142-7157,7163-7178,7184-7199,7205-7220,7231-7246,7252-7267,7273-7288,7294-7309,7315-7330,7341-7356,7362-7377,7383-7398,7404-7419,7425-7440,7451-7466,7472-7487,7493-7508,7514-7529,7535-7550,7561-7576,7582-7597,7603-7618,7624-7639,7645-7660,7671-7686,7692-7707,7713-7728,7734-7749,7755-7770,7781-7796,7802-7817,7818-7833,7839-7854,7865-7880,7891-7906,7912-7927,7933-7948,7954-7969,7970-7985,7991-8006,8012-8027,8033-8048,8054-8069,8070-8085,8091-8106,8107-8122,8128-8143,8149-8164,8170-8185,8191-8206,8212-8227,8228-8243,8249-8264,8270-8285,8291-8306,8312-8327,8328-8343,8349-8364,8370-8385,8391-8406,8412-8427,8428-8443,8449-8464,8470-8485,8491-8506,8512-8527,8528-8543,8549-8564,8570-8585,8591-8606,8607-8622,8628-8643,8649-8664,8670-8685,8686-8701,8707-8722,8728-8743,8749-8764,8770-8785,8786-8801,8807-8822,8828-8843,8849-8864,8870-8885,8891-8906,8907-8922,8928-8943,8949-8964,8970-8985,8991-9006,9007-9022,9028-9043,9049-9064,9070-9085,9091-9106,9107-9122,9128-9143,9149-9164,9170-9185,9191-9206,9207-9222,9228-9243,9249-9264,9270-9285,9291-9306,9307-9322,9328-9343,9349-9364,9370-9385,9391-9406,9407-9422,9428-9443,9449-9464,9470-9485,9491-9506,9507-9522,9528-9543,9549-956

Alignment Scores:	
Pred. No.:	5.55e-26
Length:	625
Matches:	168
Conservative:	62
Mismatches:	152
Indels:	205
Gaps:	23
DB:	1
US-09-763-153-1 (1-1245)	x TBBO (1-625)

QY	10	TTCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGATCTGTGAC	63
Db	48	PheLeuGluGluValArgLysGlyAsnLeuGluArgGluCysLeuGluGluProCysSer	67
QY	70	TTTCGAGGAGCCAGGAAATTTCCAAATGTGGATGACACACTGGCTTCTGTGCTCAAG	129
Db	68	ArgGluGluAlaPheGluAlaLeuGluSerLeuSerAlaThrAspAlaPheTrpAlaLys	87
QY	130	CAGCTCGAGCGGTGACCAGTGGTCTTGCCCTTGGAGCACCCGTGGCCACCGCTGTGC	189
Db	88	TyrThrAlaCysGluSerAla-----ArgAsnProArgGluLysLeuAsn	102
QY	190	---TGGGGCAGGCAGCTGCATCGAGGCATC-----GGCAGCTTCAGC	231
Db	103	GluCysLeuGluGlyAsnCysAlaGluGlyValGlyMetAsnTyrArgGlyAsnValSer	122
QY	232	TCCGACTGCCCCACGCGC-----TGGGAGGCGCGC-----	261
Db	123	Val---ThrArgSerGlyIleGluCysGlnLeuTrpArgSerArgTyrProHisLysPro	141
QY	262	-----TTCTGCG-----	267
Db	142	GluIleAsnSerThrHisProGlyAlaAspLeuArgGluAsnPheCysArgAsnPro	161
	268	-----CAGCGCGAG	276
QY	162	AspGlySerIleThrGlyProTrpCysTyrThrThrSerProThrLeuArgArgGluGlu	181
Db	277	GTGAGCTTCCTCAATTGCTCGGTGAC-----AACGGC	309
Db	182	CysSerValProValCysGlyGlnAspArgValThrValGluValIleProArgSerGly	201
QY	310	GGCTGCAG-----	318
Db	202	GlySerThrThrSerGlnSerProLeuLeuGluThrCysValProAspArgGlyArgGlu	221
QY	318	-----	318
Db	222	TyrArgGlyArgLeuAlaValThrThrSerGlySerArgCysLeuAlaTrpSerLeu	241
QY	319	-----CAT	321
Db	242	GlnAlaLysAlaLeuSerLysAspGlnAspPheAsnProAlaValProLeuAlaGluAsn	261
QY	322	TACTGCCTA-----GAGGAGTGGCTGGCGCGCTGTAGCTGTGCGCT	366
Db	262	PheCysArgAsnProAspGlyAspGluGluValAlaTrpCysTyrValAlaAspGlnPro	281
QY	367	GGCTTACAAGCTGGGGGACGACCTCTCGCATGTCACCCCGCAGTG-----	411
Db	282	GlyAspPheGluTyrCysAspLeuAsnTyrCysGluGluProValAspGlyAspLeuGly	301

RESULT 18  
A35827  
thrombin (EC 3.4.21.5) precursor - mouse  
C:Species: Mus musculus (house mouse)

C>Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 03-May-2002

C/Accession: A35827; A42696; S12081

R/Design, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.

DNA Cell Biol. 9, 487-498, 1990

A/Title: Characterization of the cDNA coding for mouse prothrombin and localization of

A/Reference number: A35827; MUID:91025551; PMID:222810

A/Accession: A35827

A/Status: preliminary

A/Accession: A35827

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-618 <DEG>

A/Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814

A/Experimental source: strain C57BL/6

A/Note: The data were obtained from females resulting from the cross of M. domesticus an

R/Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A/Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq

A/Reference number: A42696; MUID:92212913; PMID:1557383

A/Accession: A42696

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 384-618, 'E' <BAN>

Cross-references: GB:M81394

Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hydr

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-43/Domain: propeptide #status predicted <PRO>

F:28-88/Domain: Gla domain homology <GLA>

F:44-618/Product: prothrombin B #status predicted <MAT>

F:109-187/Domain: kringle homology <KR1>

F:215-293/Domain: kringle homology <KR2>

F:361-610/Domain: trypsin homology <TRY>

F:501-581,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status

F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,5

F:403,459,565/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.: 1.22e-25 Length: 618

Score: 533.00 Matches: 167

Percent Similarity: 39.59% Conservative: 65

Best Local Similarity: 28.50% Mismatches: 152

Query Match: 22.81% Indels: 202

DB: 2 Gaps: 27

US-09-763-153-1 (1-1245) x A35827 (1-618)

QY 1 GCCAACTCC---TTCTGGAGGAGCTCGTCACAGCAGCCTGGAGCGGAGTGCATAGAG 57

DB 44 AlaAsnSerGlyPheLeuGluGluLeuArgLysGlyAsnLeuGluArgGluCysValGlu 63

QY 58 GAGATCTGTGACTTCGAGGAGCCCAAGGAATTTCCAAATGTGGATGCACACTGGCC 117

64 GluGlnCysSerTyrGluGluAlaPheGluAlaLeuGluSerProGlnAspThrAspVal 83

118 TTCTGGTCCACACGTC-----135

DB 84 PheTrpAlaLysTyrThrValCysAspSerValArgLysProArgGluThrPheMetAsp 103

QY 135 -----135

DB 104 CysLeuGluGlyArgCysAlaMetAspLeuGlyValAsnTyrLeuGlyThrValAsnVal 123

QY 136 -----GACGGTGACACGGTGGTTCGTCCTGGAGCACCCG-----174

DB 124 ThrHisThrGlyIleGlnCysGlnLeuTrpArgSerArgTyrProHisLysProGluIle 143

QY 174 -----174

DB 144 AsnSerThrThrHisProGlyAlaAspLeuLysGluAsnPheCysArgAsnProAspSer 163

QY 175 -----TCGCC 180

DB 164 SerThrThrGlnProIleCysTyrThrThrAspProThrValArgArgGluCysSer 183



Db 525 LeuProIleValGluArgProValCysLysAlaSerThrArgIleThrAspAsn 544  
 QY 1027 ATGCTGTGTGGGGGATC---CTCGGGGAC-----CGGAGGATGCTCGGAGGGGAC 1077  
 Db 545 MetPheCysAlaGlyPheLysValAsnAspThrLysArgGlyAspAlaCysGluGlyAsp 564  
 QY 1078 ATGGGGGGCCCAVGGTC-----GCCTCCTCCACGGCACCCTGGTTCCTGGTGGGCCCTG 1131  
 Db 565 SerGlyGlyProPheValMetLysSerProPheAsnAsnArgTrpTrpGlnMetGlyIle 584  
 QY 1132 GTGAGCTGGGTGAGGGCTGTGGCTCCTTCACAACTACGGCTTTACACCAAGTCACG 1191  
 Db 595 ValSerTrpGlyGluGlyCysAspArgLysGlyLysTrpGlyPheTrpThrHisValPhe 604  
 1192 CGCTACCTCGCTGATC 1209  
 605 ArgLeuLysArgTrpIle 610

## RESULT 19

S10511  
 thrombin (EC 3.4.21.5) precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 07-May-1993 #sequence revision 07-May-1993 #text\_change 03-May-2002  
 R:Accession: S10511; A60576; B42696  
 R:DiHanch, M.; Monard, D.  
 Nucleic Acids Res. 18, 4251, 1990  
 A:Title: cDNA sequence of rat prothrombin.  
 A:Reference number: S10511; MUID:90332426; PMID:2377469  
 A:Accession: S10511  
 A:Molecule type: mRNA  
 A:Residues: 1-617 <DIR>  
 A:Cross-references: EMBL:X52835; NID:956969; PID:CAA37017.1; PID:956970  
 R:Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.  
 Endocrinology 126, 167-175, 1990  
 A:Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.  
 A:Reference number: A60576; MUID:90091942; PMID:2293980  
 A:Accession: A60576  
 A:Molecule type: protein  
 A:Residues: 44-58 <HEN>  
 A:Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat uterus.  
 R:Banfield, D.K.; MacGillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequencing.  
 A:Reference number: A42696; MUID:92212913; PMID:1557383  
 A:Accession: B42696  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 383-617, 'E' <BAN>  
 A:Cross-references: GB:M81397  
 C:Superfamily: thrombin; Gla domain homology; kringe homology; trypsin homology  
 C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolysis  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-43/Domain: propeptide #status predicted <PRO>  
 F:44-617/Product: prothrombin #status experimental <PMAT>  
 F:109-187/Domain: kringe homology <KR1>  
 F:215-292/Domain: kringe homology <KR2>  
 F:360-609/Domain: trypsin homology <TRY>  
 F:50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status  
 F:61-66,91-104,109-187,130-170,158-182,215-292,326-276,264-287,332-478,387-403,532-546,548-554,558,564/Active site: His, Asp, Ser #status predicted

## Alignment Scores:

Pred. No.:	3.57e-25	Length:	617
Score:	525.50	Matches:	165
Percent Similarity:	40.21%	Conservative:	65
Best Local Similarity:	28.85%	Mismatches:	167
Query Match:	22.45%	Indels:	175
DB:	2	Gaps:	24

US-09-763-153-1 (1-1245) x S10511 (1-617)

QY 1 GCCAACTCC---TTCTGGAGGAGCTCGTCACAGCAGCCTGGAGCGGGAGTGCATAGAG 57  
 Db 44 AlaAsnSerGlyPheLeuGluGluLeuArgLysGlyAsnLeuGluArgGluCysValGlu 63  
 QY 58 GAGATCTCTGACTTCGAGGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCC 117  
 Db 64 GluGlnCysSerTrpGluAlaPheGluAlaLeuGluSerProGlnAspThrAspVal 83  
 QY 118 TTCTGGTCCAAAGCAGCTCGACGGTGAC-----CAG 147  
 Db 84 PheTrpAlaLysTrpThrValCysAspSerValArgLysProArgGluThrPheMetAsp 103  
 QY 148 TGCTTG----- 153  
 Db 104 CysLeuGluGlyArgCysAlaMetAspLeuGlyLeuAsnTrpHisGlyAsnValSerVal 123  
 QY 154 -----GTC 156  
 Db 124 ThrHisThrGlyIleGluCysGlnLeuTrpArgSerArgTrpProHisArgProAspIle 143  
 QY 157 TTGCCCTTGAGCAGCCCGTCCGCCACCTG-----TGCTGGGGGACGGCAGC 204  
 Db 144 AsnSerThrThrHisProGlyAlaAspLeuLysGluAsnPheCysArgAsnProAspSer 163  
 QY 205 TGCATCAGCGCATCGGCAGCTTCAGCTCGAC-----TGCGGC 243  
 Db 164 SerThrSerGlyProTrpCysTrpThrAspProThrValArgArgGluGluCysSer 183  
 QY 244 -----ACGGGTGGAGGGCGCTTCTGCCAGCGAGGTGAGCTTCTCAATGTC 294  
 Db 184 IleProValCysGlyGlnGluGlyArgThrValLysMetThrProArgSerArgGly 203  
 QY 295 TCGCTGGCAAC-----GCGGCTGC----- 315  
 Db 204 SerLysGluAsnLeuSerProProLeuGlyGluCysLeuLeuGluArgGlyArgLeuTyr 223  
 QY 315 ----- 315  
 Db 224 GlnGlyAsnLeuAlaValThrThrLeuGlySerProCysLeuAlaTrpAspSerLeuPro 243  
 QY 316 -----ACGATTAC 324  
 Db 244 ThrLysThrLeuSerLysTrpGlnAsnPheAspProGluValLysLeuValGlnAsnPhe 263  
 QY 325 TGCCTA-----GAGAGGTGGCTGGCGCGCTGTAGCTGTGGCTGGC 369  
 Db 264 CysArgAsnProAspArgAspGluGluGlyAlaTrpCysPheValAlaGlnGlnProGly 283  
 QY 370 TACAAGCTGGGGACGACCTCTGCAGTGCACCCCGCAGT----- 411  
 Db 284 PheGluTyr---CysSerLeuAsnTrpCysAspGluAlaValGlyGluGluAsnHisAsp 302  
 QY 411 ----- 411  
 Db 303 GlyAspGluSerIleAlaGlyArgThrThrAspAlaGluPheHisThrPhePheAspGlu 322  
 QY 412 -----AAGTCTCCCTTGGG---AGCCCTCGAAGCGGATGAGAG 450  
 Db 323 ArgThrPheGlyLeuGlyAlaAspCysGlyLeuArgPro-----LeuPheGluLys 340  
 QY 451 AACCGGAGTCACTGAAACGACAGACAGACCAAGACCAAGTATAGATCCCGGCTC 510  
 Db 341 LysSerLeuThrAspLysThrGluLysGluLeuLeuAspSerTrpIleAspGlyArgIle 360  
 QY 511 ATGTATGGGAAGATGACACAGCGGGGACAGCCCTCGAGGTGGTCTCTCTG---GAC 567  
 Db 361 ValGluGlyTrpAspAlaGluLysGlyIleAlaProTrpGlnValMetLeuPheArgLys 380  
 QY 568 TCAAGAAGAAAGCTGGCTGGGGGAGTGTCTCATCCACCCCTCTGGGTGCTGACAGCG 627  
 Db 381 SerProGlnGluLeuCysGlyAlaSerLeuLeuSerAspArgTrpValLeuThrAla 400  
 QY 628 GCCCACTGCATG-----GATGAGTCC-----AAGAAGCTCTTGTGTC 663

Tue Mar 18 16:19:33 2003

F:202-237/Domain: fibronectin type I repeat homology <LF1>

F:245-278/Domain: EGF homology <EG2>

F:286-367/Domain: kringe homology <KR>

F:373-407/Domain: hepatocyte growth factor activator light chain status experimenta

F:408-655/Domain: hepatocyte growth factor activator heavy chain status experimenta

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

F:408-641/Domain: trypsin homology <TRY>

401 ALAHiscysileuTyrrProtrPaspLysAsnPherThrGluAsnAspLeuLeuVal 420

664 AGGCTTGAGAGATGACCTGGCGCTGGGAGAG--TGGGAGCTGGAGATC 720

421 ArgileglyLysHisSerArgThrArgTyrrGluArgAsnValGluLysSerMetIeu 440

721 AAGAGGTCTTCGTCACCCCACTACAG--AAGAGCACCCACCGACATGACATCCCA 777

441 GlulysileTyrrileHisProArGtyrAsnTrpArgGluAsnLeuAspArgAspIleAla 460

778 CTGCTGACCTGGCCGCGCCGACCTCTCCGACACATGACATGCTCCCTCCG 837

461 IeuLeuLysLysLysProValProPheSerAspTyrrileHisProValCysLeuPro 480

838 GACAGCGCCTTGACAGCGCGGAGCTCAATCAGCGCGCCGACGAGACCTCGTGTGAGCGGC 897

481 AspLysGlnThrVal--ThrSerLeuLeuGluAlaGlyTyrrLysGlyArgValThrGly 499

898 TGGGCTTACCAGCAGCGCGAG--AAGAGGCCCAAGAGAAACCCACCTTC 948

500 TrpGly-----AsnLeuArgGluThrTrpThrThrAsnIleAsnGluIleGlnProSer 517

949 GTCCTCAACTCATCAAGATCCGCTGCTCCGACCAATGAGTGACGCGAGGTCATGAGC 1008

518 ValLeuGlnValValAsnLeuProIleValGluArgProValCysLysLysAlaSerThrArg 537

1009 ACATGCTCTGAGAACATGCTGTGCGGGCATC---CTCGGGGAC-----CGGCAG 1059

538 IleArgIleThrAspAsnMetPheCysAlaGlyPheLysValAsnAspThrLysArgGly 557

1060 GATGCTGCGAGGCGACAGTGGGGGCGCATGTC-----GCCTCTTCCACGGCAC 1113

558 AspAlaCysGluGlyAspSerGlyGlyProPheValMetLysSerProTyrrAsnHisArg 577

1114 TGGTCTCTGCTGCTGTGAGCTGGGTGAGGCTGTGGGCTCTCTTCAACATACGCGC 1173

578 TrpTyrrGlnMetGlyIleValSerTrpGlyGluGlyCysAspArgAsnGlyLysTyrrGly 597

1174 GTTTACACCAAGTACGCGCTACCTCGACTGGATC 1209

598 PheTyrrThrHisValPheArgLeuLysArgTrpMet 609

RESULT 20

hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000

C:Accession: A46688

J. Biol. Chem. 268, 10024-10028, 1993

Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease

Regulation factor XII.

Reference number: A46688; MUID:93252878; PMID:7683665

Accession: A46688

Molecule type: mRNA

A:Residues: 1-655 <MIY>

A:Cross-references: DBJ:DJ14012; NID:g219680; PIDN:BAA03113.1; PID:g219681

A:Experimental source: liver (mRNA); serum (protein)

A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:131228)

A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c

C:Genetics:

A:Gene: GDB:HGFA; HGFA; HGFAP

A:Cross-references: GDB:9954514

A:Map position: 4p16-4p16

C:Function:

A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Pathway: tissue repair and regeneration

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;

C:Keywords: glycoprotein; hydrolase; kringe; liver; plasma; serine proteinase

F:1-34/Domain: signal sequence #status predicted <SIG>

F:108-148/Domain: fibronectin type II repeat homology <LF2>

F:164-197/Domain: EGF homology <EG1>

Alignment Scores: 1.01e-21 Length: 655

Pred. No.: 470.00 Matches: 138

Score: 39.84% Conservatives: 56

Percent Similarity: 28.34% Mismatches: 165

Best Local Similarity: 20.11% Indels: 128

Query Match: 1

DB: 17

US-09-763-153-1 (1-1245) x A46688 (1-655)

QY 43 CGGGAGTGCATAGAGGATCTGTGACTTCGAGAGCCCAAGGAAATTTCCAAATGTG 102

Db 195 LysaspCysGlyThrGluLysCysPheAspGluThrArgTyrrGluLeuGluGlyGly 214

QY 103 GATGACACATCGCCCTTCTGTCCCAAGCACGTCGACGCTGACCGTGCCTGTG-- 153

Db 215 AspArgTrpAlaArgValArgGlnGlyHisValGlnGlnCysGluCysPheGlyGlyArg 234

QY 154 -----GTCTGCTTGGAGCACCCTGTCGCGCGCTGCGCGCTGTCG 189

Db 235 ThrTrpCysGluGlyThrArgHisThrAlaCysLeuSerSerProCysLeuAsn-- 252

QY 190 TCGGGCAGCGCACCTGCG-----ATCGACGCGCATCGGACGCTTCAGCTGCGACTGCGCGC 243

Db 253 -----GlyGlyThrCysHisLeuIleValAlaThrGlyThrValCysAlaCysPro 270

QY 244 AGCGGTGCGGCGCGCTTCTGCGCGCGCGAGGTGAGCTTCCTCAATGCTGCTCGAC 303

Db 271 ProGlyPheAlaGlyLeuGluCysAsnIleGluProAsp---GluArgCysPheLeuGly 289

QY 304 ACGCGC----- 309

Db 290 AsnGlyThrGlyTyrrArgGlyValAlaSerThrSerAlaSerGlyLeuSerCysLeuAla 309

QY 309 ----- 309

Db 310 TrpAsnSerAspLeuLeuTyrrGlnGluLeuHisValAspSerValGlyAlaAlaLeu 329

QY 310 -----CGCTGCGACGCAATTTACTGCTTA-----GAGGAGGTGGGCTGG--- 345

Db 330 LeuGlyLeuGlyProHisAlaTyrrCysArgAsnProAspAsnAspGluArgProTrpCys 349

QY 346 -----CGGCGCTGAGCTGT 360

Db 350 TyrValValLysaspSerAlaLeuSerTrpGluTyrrCysArgLeuGluAlaCysGluSer 369

QY 361 CGCGCTGGCTACAGCTGGGGGCGACCTCTCTGAGTGTCCACCCGCGAGTGAAGTTCCCT 420

Db 370 LeuThrArgValGlnLeuSerProAspLeuLeuAlaThrLeuProGluProAlaSerPro 389

QY 421 TGTGGGAGCGCTGGAAGCGGATGGAGAGAGCGAGCGAGTCCCTCAACCTGAACACGACAGAA 480

Db 390 GlyArgGlnAlaCysGlyArgHisLysLysArgThrPheLeuArg----- 405

QY 481 GACCAAGAAGACCAAGTAGATAGTCCGCGCTCATTTGATGGAGATGACGAGCGGGGAGAC 540

Db 406 -----CAGGTGTCTCTGCTGAGCTCAAGAAGAGAGTGGCTCGCGGGGCA 594

QY 541 AGCCCTCTG----- 594

Db 419 HisProTrpLeuAlaAlaIleTyrrileGlyAspSer-----PheCysAlaGly 434

QY 595 GTGCTCATCCACCCCTCTCTGGTGTGTGACAGCGGCGGCGCTGCTGATGATGATGCTC- 648

```

Db 435 serLeuValHisThrCysTrpValValSerAlaAlaHisCysPheSerHisSerProPro 454
QY 649 ---AAGAGCTCTTGTTCAGGCTTGAGAGATGATGACCTGCGCGCTGGGAGAGTGGGAG 705
Db 455 ArgAspSerValSerValValLeuGlyGlnHisPhePheAsnArgThrThrAspValThr 474
QY 706 CTGGACCTGGACATCAAGAGGTCTTCGTCCACCCCACTACAGCAG---AGCACCAACC 762
Db 475 GlnThrPheGlyLeuGlyLysTrpThrLeuThrSerValPheAsnProSer 494
QY 763 GACATGACATCGCTGCTGCACCTGCGCCAGGCC-----GCCACCTCTCG 810
Db 495 AspHisAspLeuValLeuLeuArgLeuLysLysLysGlyAspArgCysAlaThrArgSer 514
811 CAGACCATAGTCCCTGCTCCCGCACAGCGGCTTCGAGAGCGGAGCTCAATCAG 870
515 GlnPheValGlnProIleCysLeuProGluProGly-----SerThrPhePro 530
QY 871 GCCGGCCAGAGACCTCTGTCAGCGGCTGGGGCTAC-----CAC 909
Db 531 AlaGlyHisLysCysGlnIleAlaGlyTrpGlyHisLeuAspGluAsnValSerGlyTyr 550
QY 910 AGCAGCCGAGAGAGAGGAGGCGCAAGAGAACCGCACCTTCGTCTCAACTTCATCAACATT 969
Db 551 SerSerSerLeuArgGluAla-----LeuVal 559
QY 970 CCCGTGTCCCGCACATGAGTGCAGC-----GAGGTCAATGAGCAACATGGTGTCTGAG 1023
Db 560 ProLeuValAlaAspHisLysCysSerSerProGluValTyrGlyAlaAspIleSerPro 579
QY 1024 AACATGCTGTGTGGGGCATCTCCGGGAGCCGCGAGGATCCCTGCGAGGCGACAGTGGG 1083
Db 580 AsnMetLeuCysAlaGlyTyrPheAspCysLysSerAspAlaCysGlnGlyAspSerGly 599
QY 1084 GGGCCCATGTGCTGCTCTTCCACGCGCACCCTGCTTCTGCTGGGCTGTGAGTGGGT 1143
Db 600 GlyProLeuAlaCysGluLysAsnGlyValAlaTyrLeuTyrGlyIleSerTrpGly 619
QY 1144 GAGGGCTGTGGGCTCTTCACAACTACGGCGTTTACCCAAAGTCAGCGCTACCTCGAC 1203
Db 620 AspGlyCysGlyArgLeuHisLysProGlyValTyrThrArgValAlaAsnTyrValAsp 639
QY 1204 TGGATCCATGGGCGACATCAGA 1224
Db 640 TrpIleAsnAspArgIleArg 646

```

Search completed: March 17, 2003, 17:37:13  
Job time : 70 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:01:24 ; Search time 40 Seconds  
(without alignments)  
2581.901 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=framed\_n2p\_model -DEV=xlp  
-O=Csqn2\_1/USPTO\_spool/US09763153/runat\_11032003\_084247\_1981/app\_query.fasta\_1.1415  
-DB=SwissProt\_40 -QMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0  
-LIST=60 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=ptc -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-USER=US09763153 @Csqn2\_1.16.runat\_11032003\_084247\_1981 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	2298	98.3	461	1 PRTC_HUMAN	P04070 homo sapien
2	1872.5	80.1	458	1 PRTC_RABIT	Q28661 oryctolagus
3	1750.5	74.9	459	1 PRTC_PIG	Q981p2 sus scrofa
4	1660	71.0	456	1 PRTC_BOVIN	P00745 bos taurus
5	1650.5	70.6	461	1 PRTC_RAT	P31394 rattus norv
6	1637.5	70.1	461	1 PRTC_MOUSE	P33587 mus musculus
7	835	35.7	161	1 PRTC_MACMU	Q28506 macaca mula
8	814.5	34.9	490	1 PRTC_HUMAN	O19045 oryctolagus
9	809	34.6	492	1 FA10_BOVIN	P00743 bos taurus
10	802	34.3	444	1 FA7_RABIT	P98139 oryctolagus
11	801.5	34.3	475	1 FA10_CHICK	P25155 gallus gall
12	801.5	34.3	488	1 FA10_HUMAN	P00742 homo sapien
13	783	33.5	466	1 FA7_HUMAN	P08709 homo sapien
14	779.5	33.4	407	1 FA7_BOVIN	P22457 bos taurus
15	770	32.9	446	1 FA7_MOUSE	P70375 mus musculus
16	763	32.6	452	1 FA9_CANFA	P19540 canis famil
17	736	31.5	461	1 FA9_HUMAN	P00740 homo sapien
18	726	31.1	459	1 FA9_MOUSE	P16294 mus musculus

19	724	31.0	376	1	FA10_TROCA	P81428 tropidechis
20	717	30.7	157	1	PRTC_CANFA	Q28278 canis famil
21	716	30.6	157	1	PRTC_FELCA	Q28412 felis silve
22	714.5	30.6	416	1	FA9_BOVIN	P00741 bos taurus
23	700	30.0	157	1	PRTC_HORSE	Q28380 equus cabal
24	661	28.3	157	1	PRTC_CAPHI	Q28315 capra hircu
25	562.5	24.1	622	1	THRB_HUMAN	P00734 homo sapien
26	538.5	23.0	625	1	THRB_BOVIN	P00735 bos taurus
27	533	22.8	618	1	THRB_MOUSE	P18221 mus musculus
28	525.5	22.5	617	1	THRB_RAT	P18292 rattus norv
29	475.5	20.3	653	1	HGFA_MOUSE	Q9R098 mus musculus
30	470	20.1	655	1	HGFA_HUMAN	Q04756 homo sapien
31	468	20.0	400	1	PRT2_HUMAN	P22891 homo sapien
32	445.5	19.1	396	1	PRT2_BOVIN	P00744 bos taurus
33	439.5	18.8	275	1	TRYT_PIG	Q9N2d1 sus scrofa
34	427	18.3	271	1	FA9_PIG	P16293 sus scrofa
35	426	18.2	638	1	KAL_MOUSE	P26262 mus musculus
36	424.5	18.2	638	1	KAL_HUMAN	P03952 homo sapien
37	421	18.0	455	1	TMS5_MOUSE	Q9er04 mus musculus
38	420.5	18.0	699	1	CRAR_HUMAN	P48740 h complemen
39	420	18.0	275	1	FA9_RABIT	P16292 oryctolagus
40	419.5	18.0	812	1	PLMN_MOUSE	P20918 mus musculus
41	419	17.9	625	1	FA11_HUMAN	P03951 homo sapien
42	418	17.9	490	1	TMS2_MOUSE	Q9j1q8 mus musculus
43	417.5	17.9	855	1	ST14_MOUSE	P56677 mus musculus
44	417	17.8	285	1	FA9_CAVPO	P16295 cavia porce
45	415.5	17.8	704	1	CRAR_MOUSE	P98064 mus musculus
46	414	17.7	638	1	KAL_RAT	P14272 rattus norv
47	411	17.6	786	1	STUB_DROME	Q05319 drosophila
48	411	17.6	1069	1	ENTK_MOUSE	P97435 mus musculus
49	410	17.5	810	1	PLMN_HUMAN	P00747 homo sapien
50	409.5	17.5	282	1	FA9_RAT	P16296 rattus norv
51	408.5	17.5	875	1	NETR_HUMAN	P56730 homo sapien
52	406	17.4	251	1	KLKE_HUMAN	Q9P0g3 homo sapien
53	406	17.4	417	1	HSPS_HUMAN	P05981 homo sapien
54	405.5	17.4	492	1	TMS2_HUMAN	O15393 homo sapien
55	404.5	17.3	275	1	TRB1_HUMAN	O15661 homo sapien
56	404	17.3	1019	1	LFC_TACTR	P28175 tachypleus
57	403.5	17.3	275	1	TRB2_HUMAN	P20231 homo sapien
58	403	17.2	273	1	TRYT_SHEEP	Q9Xsm2 ovis aries
59	403	17.2	583	1	CFAL_HUMAN	P05156 homo sapien
60	402.5	17.2	275	1	TRYA_HUMAN	P15157 homo sapien

## ALIGNMENTS

RESULT 1  
PRTC\_HUMAN  
ID PRTC\_HUMAN STANDARD; PRT: 461 AA.  
AC P04070; Q16001; Q15190; Q15189;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 01-NOV-1986 (Rel. 03, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).  
DE PROC.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=85270390; PubMed=2991887;  
RA Foster D.C., Yoshitake S., Davie E.W.;  
RT "The nucleotide sequence of the gene for human protein C.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=85269639; PubMed=2991859;  
RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutsky J., Crabtree G.R.,  
RA Long G.L.;  
RT "The structure and evolution of a 461 amino acid human protein C

precursor and its messenger RNA, based upon the DNA sequence of  
cloned human liver cDNAs.";  
Nucleic Acids Res. 13:5233-5247(1985).  
[3]  
SEQUENCE FROM N.A.  
MEDLINE-86120978; PubMed-3511471;  
Plutsky J., Hoskins J.A., Long G.L., Crabtree G.R.;  
Evolution and organization of the human protein C gene.";  
Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).  
[4]  
SEQUENCE FROM N.A.  
Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
Rickerson D.A.;  
Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE OF 106-461 FROM N.A.  
MEDLINE-84272714; PubMed-6589623;  
Foster D.C., Davie E.W.;  
Characterization of a cDNA coding for human protein C.";  
Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).  
[6]  
CARBOHYDRATE-LINKAGE SITE ASN-371.  
MEDLINE-90293094; PubMed-1694179;  
Mileich J.P., Broze G.J. Jr.;  
Beta protein C is not glycosylated at asparagine 329. The rate of  
translation may influence the frequency of usage at asparagine-X-  
cysteine sites.";  
J. Biol. Chem. 265:11397-11404(1990).  
[7]  
HYDROXYLATION.  
MEDLINE-92184750; PubMed-1544894;  
Harris R.J., Ling V.T., Spellman M.W.;  
O-linked fucose is present in the first epidermal growth factor  
domain of factor XII but not protein C.";  
J. Biol. Chem. 267:5102-5107(1992).  
[8]  
3D-STRUCTURE MODELING OF 175-450.  
MEDLINE-94272342; PubMed-8003577;  
Fisher C.L., Greengard J.S., Griffin J.H.;  
Models of the serine protease domain of the human antithrombotic  
plasma factor activated protein C and its zymogen.";  
Protein Sci. 3:588-599(1994).  
[9]  
X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.  
MEDLINE-97157472; PubMed-9003757;  
Mather I., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,  
Rode W.;  
The 2.8 A crystal structure of Gla-domainless activated protein C.";  
EMBO J. 15:6822-6831(1996).  
[10]  
REVIEW ON PROC VARIANTS.  
MEDLINE-93190290; PubMed-8446940;  
Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,  
Sala N., Cooper D.N.;  
Protein C deficiency: a database of mutations. For the Protein C & S  
Subcommittee of the Scientific and Standardization Committee of the  
International Society on Thrombosis and Haemostasis.";  
Thromb. Haemost. 69:77-84(1993).  
[11]  
VARIANT CYS-444.  
MEDLINE-87204221; PubMed-2437584;  
Romeo G., Hassan H.J., Staepfli S., Roncuizi L., Cianetti L.,  
Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,  
Cortese R.;  
Hereditary thrombophilia: identification of nonsense and missense  
mutations in the protein C gene.";  
Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).  
[12]  
VARIANT TRP-211 (LONDON-1).  
MEDLINE-90098906; PubMed-2602169;  
Grundt C.B., Chitolie A., Talbot S., Bevan D., Kakkar V.V.,  
Cooper D.N.;  
Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in  
the protein C gene causing thrombosis.";  
Nucleic Acids Res. 17:10513-10513(1989).  
[13]  
VARIANT CYS-272.  
MEDLINE-91329836; PubMed-1868249;  
Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;  
The spectrum of genetic defects in a panel of 40 Dutch families with  
symptomatic protein C deficiency type I: heterogeneity and founder  
effects.";  
Blood 78:890-894(1991).  
[14]  
VARIANTS ALA-62 (VERMONT-1) AND MET-76.  
MEDLINE-92190481; PubMed-1347706;  
Bavali E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,  
Rainville I.R., Long G.L.;  
Protein C Vermont: symptomatic type II protein C deficiency  
associated with two GLA domain mutations.";  
Blood 79:1456-1465(1992).  
[15]  
VARIANT ASP-418 (HONG KONG-2).  
MEDLINE-92305321; PubMed-1611081;  
Sugahara Y., Miura O., Yuen P., Aoki N.;  
Protein C deficiency Hong Kong 1 and 2: hereditary protein C  
deficiency caused by two mutant alleles, a 5-nucleotide deletion and  
a missense mutation.";  
Blood 80:126-133(1992).  
[16]  
VARIANT LEU-289.  
MEDLINE-92380660; PubMed-1511988;  
Grundt C.B., Chisholm M., Kakkar V.V., Cooper D.N.;  
A novel homozygous missense mutation in the protein C (PROC) gene  
causing recurrent venous thrombosis.";  
Hum. Genet. 89:683-684(1992).  
[17]  
VARIANTS GLN-220 AND TRP-220.  
MEDLINE-92380661; PubMed-1511989;  
Grundt C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;  
Two different missense mutations at Arg 178 of the protein C (PROC)  
gene causing recurrent venous thrombosis.";  
Hum. Genet. 89:685-686(1992).  
[18]  
VARIANT GLN-220.  
MEDLINE-93250852; PubMed-1301959;  
Gandrille S., Vidaud M., Aiach M., Alhenc-Gelas M., Fischer A.M.,  
Cavaud-Hellman M., Toulon P., Flessinger J.N., Goossens M.;  
Two novel mutations responsible for hereditary type I protein C  
deficiency: characterization by denaturing gradient gel  
electrophoresis.";  
Hum. Mutat. 1:491-500(1992).  
[19]  
VARIANT SER-334.  
MEDLINE-92276939; PubMed-1593215;  
Yamamoto K., Matsushita T., Suglura I., Takamatsu J., Iwasaki E.,  
Wada H., Deguchi K., Shirakawa S., Saio H.;  
Homozygous protein C deficiency: identification of a novel missense  
mutation that causes impaired secretion of the mutant protein C.";  
J. Lab. Clin. Med. 119:682-689(1992).  
[20]  
VARIANTS TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.  
MEDLINE-93313192; PubMed-8324221;  
Gandrille S., Alhenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,  
Juhani-Vague I., Aiach M.;  
Five novel mutations located in exons III and IX of the protein C  
gene in patients presenting with defective protein C anticoagulant  
activity.";  
Blood 82:159-168(1993).  
[21]  
VARIANTS G-14; Q-211; Y-244; Q-253; L-321; C-328; I-385; T-388 AND  
V-388.  
MEDLINE-93271391; PubMed-8499565;  
Poort S.R., Fabinger-Fasching I., Mannhalter C., Reitsma P.H.,  
Bertina R.M.;  
Twelve novel and two recurrent mutations in 14 Austrian families



with hereditary protein C deficiency.";  
 RL Blood Coagul. Fibrinolysis 4:273-280(1993).  
 [22]  
 RP VARIANT TRP-57  
 RX MEDLINE=93271396; PubMed=8499568;  
 RA Miller D.S., Grundy C.B., Bignell P., Moffat E.H., Martin R.,  
 RA Kakkar V.V., Cooper D.N.;  
 RT "A Gla domain mutation (Arg 15-->Trp) in the protein C (PROC) gene  
 RT causing type 2 protein C deficiency and recurrent venous  
 RT thrombosis";  
 RL Blood Coagul. Fibrinolysis 4:345-347(1993).  
 RN [23]  
 RP VARIANTS R-145; L-210; W-211; T-243; L-321; M-340 AND Y-426.  
 MEDLINE=94122329; PubMed=8292730;  
 RX Tsay W., Greengard J.S., Montgomery R.R., McPherson R.A., Fucci J.C.,  
 RA Koepfer M.A., Coughlin J., Griffin J.H.;  
 RT "Genetic mutations in ten unrelated American patients with  
 RT symptomatic type 1 protein C deficiency.";  
 RL Blood Coagul. Fibrinolysis 4:791-796(1993).  
 RN [24]  
 RP VARIANT SER-423.  
 RX MEDLINE=94001606; PubMed=8398832;  
 RA Marchetti G., Patrachini P., Gemmati D., Castaman G., Rodeghiero F.,  
 RA Wacey A., Cooper D.N., Tuddenham E.G., Bernardi F.;  
 RT "Symptomatic type II protein C deficiency caused by a missense  
 RT mutation (Gly 381-->Ser) in the substrate-binding pocket.";  
 RL Br. J. Haematol. 84:285-289(1993).  
 RN [25]  
 RP SEQUENCE OF 43-64 FROM N.A., AND VARIANT GLY-57 (YONAGO).

## Alignment Scores:

Pred. No.: 1,41e-114 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x PRTC\_HUMAN (1-461)

QY 1 GCCAACTCTTCCTGGAGAGCTCGTCACAGCAGCTGGAGCGGAGTCATAGAGAG 60  
 DB 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysileGluGlu 62  
 QY 61 ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAATATGTGATGACACTGGCCTTC 120  
 DB 63 IleCysAspPheGluGluAlaLysGluilePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGGTCCCAAGCAGTCGACGGTGACCACTGTGCTGTGCTGCTGGAGACCCGTCGCC 180  
 DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCCTGTGCTCGGGGACGGCAGCTGCATCGAGCAGCTTCAGCTCGAGCTGC 240  
 DB 103 SerLeuCysCysGlyHisGlyThrCysileAspGlyileGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGAGGCGCGCTTCGCCAGCGGAGTGAGCTTCCTCAATGCTCGCTG 300  
 DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAACGGCGCTCGACGCTACTGCTAGAGAGTGCGCTGGCGGCGCTGTAGCTGT 360  
 DB 143 AspasnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgCysSerCys 162  
 QY 361 GCGCTGGCTACAGCTGGGGGACGACCTCTGCTAGCTGTCCACCCCGCAGTGAAGTCCCT 420  
 DB 163 AlaProGlyTyrcysLeuGlyAspAspLeuGluGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAAAGCGAGTCACCTGAAACGAGACAGAA 480  
 DB 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAACCAAGTAGATCCGCGCTCATTCATGGGGAAGATGACAGCGGGGAGAC 540

DB 203 AspGlnGluAspGlnValAspProArgLeuileAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGCAGGTGGTCTCTGCTGACTCAAGAGAGAGCTGGCTGGGGGGAGTGCCTC 600  
 DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTCTGGTGTCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGAT 660  
 DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
 QY 661 GTCAGGCTGGAGAGTATGACCTGGGGCTGGGAGAGTGGAGTGGAGTGGAGTGGAGTGG 720  
 DB 263 ValArgLeuGlyGlyLysTrpAspLeuArgArgTrpLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGAGGTCTCTGCTCCACCCCACTACAGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAG 780  
 DB 283 LysGluValPheValHisProAsnTyrcysSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGACCTGGCCCGCCAGCCCGCCCTCTCGCAGACCATAGTGCCTCTCTCTCTCTCTCT 840  
 DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGCCCTTCGAGAGCGGAGCTCAATCAGCCCGCCAGAGACCTCTCTGTCAGCGGCTGG 900  
 DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACCCACAGCAGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 DB 343 GlyThrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCTGGTTCCTCGCCACATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 363 IleLysileProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGTGTGCTGGGCGATCTCTCGGGGACCGGAGGAGGAGGAGGAGGAGGAGG 1080  
 DB 383 GluAsnMetLeuGlyAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCCCCATGTGCTGCT 1140  
 DB 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCT 1200  
 DB 423 GlyGlyGlyCysGlyLeuLeuHisAsnTyrcysGlyValTyrcysValSerArgTyrcys 442  
 QY 1201 GACTGGATCCATGGGCGACATCAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245  
 DB 443 AspTrpIleHisGlyHisileArgAspLysGluAlaProGlnLys 457  
 RESULT 2  
 PRTC\_RABIT  
 ID PRTC\_RABIT STANDARD; PRT; 458 AA.  
 AC Q28661;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
 DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation  
 DE factor XIV) (Fragment).  
 GN PROC.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Shen L., He X., Dahlback B.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA







CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85014826; PubMed=6091100;  
 RA Long G.L., Balagaje R.M., McGillivray R.T.A.;  
 RT "Cloning and sequencing of liver cDNA coding for bovine protein C.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).  
 RN [2]  
 RP SEQUENCE OF 40-194.  
 RX MEDLINE=83007325; PubMed=6896876;  
 RA Fernlund P., Stenflo J.;  
 RT "Amino acid sequence of the light chain of bovine protein C.";  
 RL J. Biol. Chem. 257:12170-12179(1982).  
 RN [3]  
 RP REVISION TO 110.  
 RX MEDLINE=83169769; PubMed=6572939;  
 RA Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;  
 RT "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).  
 RN [4]  
 RP SEQUENCE OF 197-456.  
 RX MEDLINE=83007326; PubMed=6896877;  
 RA Stenflo J., Fernlund P.;  
 RT "Amino acid sequence of the heavy chain of bovine protein C.";  
 RL J. Biol. Chem. 257:12180-12190(1982).  
 RN [5]  
 RP PROCESSING, AND CALCIUM-BINDING DATA.  
 RX MEDLINE=83213513; PubMed=6304092;  
 RA Esmon N.L., Debaul L.E., Esmon C.T.;  
 RT "Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless protein C.";  
 RL J. Biol. Chem. 258:5548-5553(1983).  
 RN [6]  
 RP PROCESSING, AND CALCIUM-BINDING DATA.  
 RX MEDLINE=83213514; PubMed=6406503;  
 RA Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;  
 RT "Structural changes required for activation of protein C are induced by Ca2+ binding to a high affinity site that does not contain gamma-carboxyglutamic acid";  
 RL J. Biol. Chem. 258:5554-5560(1983).  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
 CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va and VIIIA.  
 CC -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A TETRAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS STRONGLY PROMOTED BY THROMBOMODULIN.  
 CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
 CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
 CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING SITE IS NECESSARY FOR THE RECOGNITION OF THE THROMBIN-THROMBOMODULIN COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; K02435; AAA30685.1; -.

DR PIR; A00928; KXBO.  
 DR HSSP; P04070; LPCU.  
 DR MEROPS; S01.218; -.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00594; gla; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00065; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_Ca; 1.  
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Blood coagulation; Glycoprotein; Serine protease;  
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 29  
 FT PROPEP 30 39  
 FT CHAIN 40 194  
 FT CHAIN 197 456  
 FT PEPTIDE 197 210  
 FT DOMAIN 94 129  
 FT DOMAIN 133 173  
 FT DOMAIN 211 456  
 FT MOD\_RES 45 45  
 FT MOD\_RES 46 46  
 FT MOD\_RES 53 53  
 FT MOD\_RES 55 55  
 FT MOD\_RES 58 58  
 FT MOD\_RES 59 59  
 FT MOD\_RES 62 62  
 FT MOD\_RES 64 64  
 FT MOD\_RES 65 65  
 FT MOD\_RES 68 68  
 FT MOD\_RES 74 74  
 FT MOD\_RES 110 110  
 FT ACT\_SITE 252 252  
 FT ACT\_SITE 298 298  
 FT ACT\_SITE 397 397  
 FT DISULFID 56 61  
 FT DISULFID 89 108  
 FT DISULFID 98 103  
 FT DISULFID 102 117  
 FT DISULFID 119 128  
 FT DISULFID 137 148  
 FT DISULFID 144 157  
 FT DISULFID 159 172  
 FT DISULFID 180 318  
 FT DISULFID 237 253  
 FT DISULFID 368 383  
 FT DISULFID 393 421  
 FT CARBOHYD 136 136  
 FT CARBOHYD 289 289  
 FT CARBOHYD 350 350  
 FT CARBOHYD 366 366  
 FT VARIANT 82 82  
 FT CONFLICT 455 456  
 SQ SEQUENCE 456 AA; 51407 MW; CAAF683F894C209 CRC64;  
 VP -> PV (IN REF. 4).  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 F -> K.  
 Length: 456  
 Matches: 298  
 Conservative: 38

Alignment Scores:  
 Pred. No.: 6.74e-81  
 Score: 1660.00  
 Percent Similarity: 80.96%

Best Local Similarity: 71.81% Mismatches: 73  
Query Match: 71.03% Indels: 6  
DB: 1 Gaps: 2

US-09-763-153-1 (1-1245) x PRTC\_BOVIN (1-456)

QY 1 GCCAACTCTCTCTGGAGGAGTCCCTGCACAGCGCTCGAGCGGAGTGATAGAGGAG 60  
Db 40 ALaAsnSerPheLeuGluGluLeuArgProGlyAsnValGluArgGluCysSerGluGlu 59  
QY 61 ATCTGTCACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCTTC 120  
Db 60 ValCysGluPheGluGluAlaArgGluLeuPheGlnAsnThrGluAspThrMetAlaPhe 79  
QY 121 TGGTCCAGACGCTGACGGTGACAGTGTCTGCTGCTTCCCTTGGAGCACCGTCGCGC 180  
Db 80 TrpSerPheTyrSerAspGlyAspGlnCysGluAspArgProSerGlySerProCysAsp 99  
QY 181 AGCTGTGCTCGCGGACGCGCTGCATCGACGCGATCGGAGCTTCAGCTCGCACTGC 240  
Db 100 LeuProCysCysGlyArgGlyLysCysIleAspGlyLeuGlyGlyPheArgCysAspCys 119  
QY 241 CGCAGCGGCTGGAGGCGGCTTCCTGCACGCGGAGGTGAGCTTCCTCAATGCTCGCTG 300  
Db 120 AlaGluGlyTrpGluGlyArgPheCysLeuHisGluValArgPheSerAsnCysSerAla 139  
QY 301 GACAAACGCGCTGCACGCTACTCTAGAGGAGGTGGCTGCGGCGCTGTAGCTGT 360  
Db 140 GluAsnGlyGlyCysAlaHisTyrCysMetGluGluGluArgHisCysSerCys 159  
QY 361 GCGCTGCTTACAGCTGGGAGGACGCTCTCTGAGTGTACCCCGAGTGAAGTTCCTT 420  
Db 160 AlaProGlyTyrArgLeuGluAspAspHisGlnLeuCysValSerLysValThrPhePro 179  
QY 421 TGTGGAGGCGCTGGAGCGATGGAGAAAGCGAGTCACCTGAAACGAGACACAGAA 480  
Db 180 CysGlyArgLeuGlyLysArgMetGluLysArgLysThrLeuLysArgAspThrAsn 199  
QY 481 -----GACCAAGAACACCAAGTAGATCCGCGCTCATGTAGGAGAGTACCAAGAG 534  
Db 200 GlnValAspGlnLysAspGlnLeuAspProArgIleValAspGlyGlnGluAlaGlyTrp 219  
QY 535 GGACACGCGCTCGGAGGTGGCTCTGCTGACTCAAGAGAGCTGGCTCGCGGCGCA 594  
Db 220 GlyGluSerProThrGlnAlaValLeuLeuAspSerLysLysLeuValCysGlyAla 239  
QY 595 GTGCTCATCCACCCTCTCGGTGCTGACAGCGCGCCACTGCATGATGATGATGATGATG 654  
Db 240 ValLeuIleHisValSerTrpValLeuThrValAlaHisCysLeuAspSerArgLysLys 259  
QY 655 CTCCTGTCAGGCTGGAGAGTATGACTCGCGGCTGGAGAGTGGAGCTGGAGCTG 714  
Db 260 LeuIleValArgLeuGlyGluTyrAspMetArgArgTrpGluSerTrpGluValAspLeu 279  
QY 715 GACATCAAGAGGTCTTCGCTCCACCCCACTACAGCAAGAGCACCACCGACATGACATC 774  
Db 280 AspIleLysGluValIleHisProAsnTyrThrLysSerThrSerAspAsnAspIle 299  
QY 775 GCATGCTGACCTCGCGGCGCGCCGCTCTCGCAGACCATAGTGCCTCTGCTC 834  
Db 300 AlaLeuLeuArgLeuAlaLysProAlaThrLeuSerGlnThrIleValProIleCysLeu 319  
QY 835 CCGACAGCGGCTTCGAGGCGGAGCTCAATCAGCGCGCGGAGGAGGAGGAGGAGGAGG 894  
Db 320 ProAspSerGlyLeuSerGluArgLysLeuThrGlnValGlnGluThrValValThr 339  
QY 895 GGCTGGGCTACACAGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 954  
Db 340 GlyTrpGlyTyrArgAsp-----GluThrLysArgAsnArgThrPheValLeu 355  
QY 955 AACTTCATCAAGATTCGCTGGTCCGCGACCAATGAGTGCAGCGGAGGTTCATGACCAATG 1014  
Db 356 SerPheIleLysValProValProValProTyrAsnAlaCysValHisAlaMetGluAsnLys 375

QY 1015 GTGCTGTGAGAACATGCTGTGTGCGGCATCTCGGGACCGACGAGATGCTCGAGGCG 1074  
Db 376 IleSerGluAsnMetLeuCysAlaGlyIleLeuGlyAspProArgAspAlaCysGluGly 395  
QY 1075 GACAGTGGGGGCCCATGGTGCCTCTCCACGCGACCTGGTTCCTGGTGGGCGCTGGTG 1134  
Db 396 AspSerGlyGlyProMetValThrPhePheArgGlyThrTrpPheLeuValGlyLeuVal 415  
QY 1135 AGCTGGGTGAGGCTGTGGGCTCTTACAACTACGCGGTTTACACCAAGTCCAGCGC 1194  
Db 416 SerTrpGlyGluGlyCysGlyArgLeuTyrAsnTyrGlyValTyrThrLysValSerArg 435  
QY 1195 TACCTCGACTGCATCCATGGCGCACATCAGACAGCAAGGAGGCGCC 1239  
Db 436 TyrLeuAspTrpIleTyrGlyHisIleLysAlaGlnGluAlaPro 450

RESULT 5  
PRTC\_RAT  
ID PRTC\_RAT STANDARD; PRT; 461 AA.  
AC P31394;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
DE (Autoproteolysin II) (Anticoagulant protein C) (Blood coagulation factor XIV).  
GN PROC.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Wistar; TISSUE=Liver;  
RX MEDLINE=92329550; PubMed=1627650;  
RA Okafuji T., Maekawa K., Nawa K., Marumoto Y.;  
RT "The cDNA cloning and mRNA expression of rat protein C";  
RL Blochm. Biophys. Acta 1131:329-332(1992).  
CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
CC and VIIIA.  
CC -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS  
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
CC STRONGLY PROMOTED BY THROMBOMODULIN.  
CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
CC SITE IS NECESSARY FOR THE RECOGNITION OF THE  
CC THROMBIN-THROMBOMODULIN COMPLEX.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X64336; CAA45617.1; -.  
CC PIR; S18994; S18994.  
CC PIR; S24312; S24312.  
CC HSP; P04070; LPCU.  
CC MEROPS; S01.218; -.



DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001254; Ser.protease\_Try.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_Ca; 1.  
 DR SMART: SM00001; EGF-like; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01187; EGF\_Ca; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 DR Blood coagulation; Glycoprotein; Serine protease;  
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
 FT SIGNAL 1 32  
 FT PROPEP 33 41  
 FT CHAIN 42 196  
 FT CHAIN 199 461  
 FT PEPTIDE 199 212  
 FT SITE 212 213  
 FT DOMAIN 96 131  
 FT DOMAIN 135 175  
 FT DOMAIN 213 461  
 FT MOD\_RES 47 47  
 FT MOD\_RES 48 48  
 FT MOD\_RES 55 55  
 FT MOD\_RES 57 57  
 FT MOD\_RES 60 60  
 FT MOD\_RES 61 61  
 FT MOD\_RES 66 66  
 FT MOD\_RES 67 67  
 FT MOD\_RES 70 70  
 FT MOD\_RES 112 112  
 FT ACT\_SITE 254 254  
 FT ACT\_SITE 300 300  
 FT ACT\_SITE 402 402  
 FT DISULFID 58 63  
 FT DISULFID 91 110  
 FT DISULFID 100 105  
 FT DISULFID 104 119  
 FT DISULFID 121 130  
 FT DISULFID 139 150  
 FT DISULFID 146 159  
 FT DISULFID 161 174  
 FT DISULFID 182 320  
 FT DISULFID 239 255  
 FT DISULFID 373 387  
 FT DISULFID 398 426  
 FT CARBOHYD 215 291  
 FT CARBOHYD 291 355  
 FT CARBOHYD 355 355

InterPro: IPR000152; Asx\_hydroxyl.  
 InterPro: IPR001314; Chymotrypsin.  
 InterPro: IPR000561; EGF-like.  
 InterPro: IPR001881; EGF\_Ca.  
 InterPro: IPR002383; GLA\_blood.  
 InterPro: IPR001254; Ser.protease\_Try.  
 InterPro: IPR000294; VitK\_dep\_GLA.  
 Pfam: PF00008; EGF; 2.  
 Pfam: PF00089; trypsin; 1.  
 Pfam: PF00594; gla; 1.  
 PRINTS: PR00722; CHYMOTRYPSIN.  
 PRINTS: PR00001; GLABLOOD.  
 SMART: SM00179; EGF\_Ca; 1.  
 SMART: SM00001; EGF-like; 1.  
 SMART: SM00069; GLA; 1.  
 SMART: SM00020; Tryp\_Spc; 1.  
 PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 PROSITE: PS00022; EGF\_1; 1.  
 PROSITE: PS01186; EGF\_2; 2.  
 PROSITE: PS01187; EGF\_Ca; 1.  
 PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 PROSITE: PS02040; TRYPSIN\_DOM; 1.  
 PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 PROSITE: PS00135; TRYPSIN\_SER; 1.  
 Blood coagulation; Glycoprotein; Serine protease;  
 Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
 SIGNAL 1 32  
 PROPEP 33 41  
 CHAIN 42 196  
 CHAIN 199 461  
 PEPTIDE 199 212  
 SITE 212 213  
 DOMAIN 96 131  
 DOMAIN 135 175  
 DOMAIN 213 461  
 MOD\_RES 47 47  
 MOD\_RES 48 48  
 MOD\_RES 55 55  
 MOD\_RES 57 57  
 MOD\_RES 60 60  
 MOD\_RES 61 61  
 MOD\_RES 66 66  
 MOD\_RES 67 67  
 MOD\_RES 70 70  
 MOD\_RES 112 112  
 ACT\_SITE 254 254  
 ACT\_SITE 300 300  
 ACT\_SITE 402 402  
 DISULFID 58 63  
 DISULFID 91 110  
 DISULFID 100 105  
 DISULFID 104 119  
 DISULFID 121 130  
 DISULFID 139 150  
 DISULFID 146 159  
 DISULFID 161 174  
 DISULFID 182 320  
 DISULFID 239 255  
 DISULFID 373 387  
 DISULFID 398 426  
 CARBOHYD 215 291  
 CARBOHYD 291 355  
 CARBOHYD 355 355

SQ SEQUENCE 461 AA; 51912 MW; 884CF93664EDACD5 CRC64;  
 Alignment Scores:  
 Pred. No.: 2.14e-80 Length: 461  
 Score: 1650.50 Matches: 289  
 Percent Similarity: 82.73% Conservative: 56  
 Best Local Similarity: 69.30% Mismatches: 69  
 Query Match: 70.62% Indels: 3  
 DB: 1 Gaps: 2  
 US-09-763-153-1 (1-1245) x PRTC\_RAT (1-461)  
 QY 1 GCCAACTCCTCTCGAGGAGCTCGTCCACAGCACGCTGGAGCGAGTGCATAGAGGAG 60  
 DB 42 AlaasnSerPheLeuGluValArgAlaGlySerLeuGluArgGlyCysMetGluGlu 61  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACTGGCCCTC 120  
 DB 62 IleCysAspPheGluAlaGlnGluPheGlnAsnValGluAspThrLeuAlaPhe 81  
 QY 121 TGGTCCAAAGCAGTCGACGGTGACCGTGTGGTCTTGGCTTGGAGACACCGGTGCC 180  
 DB 82 TrpIleLysTyrPheAspGlyAspGlnCysSerThrProLeuAspHisGlnCysAsp 101  
 QY 181 AGCCTGTGCTGGGGCAGCGGACGTCATCGACGCGCATCGGACGCTTCAGCTGCCACTGC 240  
 DB 102 SerProCysCysGlyHisGlyThrCysIleAspGlyLeuGlyPheSerCysSerCys 121  
 QY 241 CGCAGCGCTGGAGGCGGCTTCTGCGAGCGGAGGTGAGTCTCCTCAATGTGCTGCTG 300  
 DB 122 AsplysGlyTyrGluGlyArgPheCysGlnGlnGluMetGlyPheGlnAspCysArgVal 141  
 QY 301 GACAAAGCGGCTGACCGCATCTACTGCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360  
 DB 142 LysAsnGlyGlyCysTyrHisThrCysLeuGluGluThrArgGlyArgCysArgCys 161  
 QY 361 GCGCTGCTGCTACAGCTGGGGAGGACCTCTGCGAGTGTACCCGCGAGTCAAGTTCGCT 420  
 DB 162 AlaProGlyTyrGluLeuAlaAspHisMetHisCysArgProThrValAsnPhePro 181  
 QY 421 TGTGGAGGCGCTGGAAGCGGATGGAGAAGCGGAGTACCTGACCTGAAACGA-----GAC 474  
 DB 182 CysGlyLysLeuTyrLysArgThrAspLysArgLysAsnPheLysArgAspIleAsp 201  
 QY 475 ACAGAAGACCAAGAACCAAGTAGTCCGCGCTCATGTATGATGGAGAGATGACAGCGCG 534  
 DB 202 ProGluAspGluGluLeuGluGlyProArgIleValAsnGlyThrLeuThrLysGln 221  
 QY 535 GGAGACAGCCCTCGCAGGTGCTCTGCTGAGCTCAAGAGAGAGCTGGCTGCGGGGCA 594  
 DB 222 GlyAspSerProTyrGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGly 241  
 QY 595 GTGCTCATCCACCCTCTCTGGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGAT 654  
 DB 242 ValLeuIleHisThrSerTyrValLeuThrAlaAlaHisCysLeuGluSerSerLysLys 261  
 QY 655 CTCCTTGTGAGCTGTGAGAGTATGACCTGCGGCGCTGGGAGAGAGTGGAGCTGGACCTG 714  
 DB 262 LeuThrValArgLeuGlyGlyTyrAspLeuArgArgAspProTyrGluLeuAspLeu 281  
 QY 715 GACATCAAGGAGGTCTTCTGCCACCCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 774  
 DB 282 AspIleLysGluValLeuValHisProAsnTyrThrArgSerAsnSerAspAspIle 301  
 QY 775 GCACTGTGCACTGGGCCCGCCGCGCCCTCTCGAGAGACCATAGTGGCCATCTGCTCCT 834  
 DB 302 AlaLeuLeuArgLeuSerGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 321  
 QY 835 CCGGACAGCGGCTTGGAGAGCGGAGCTCAATCAGCGCGCGGAGAGACCTCTGTGAGC 894  
 DB 322 ProAsnSerGlyLeuAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340  
 QY 895 GGCTGGGCTTACCACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954



FT	MOD_RES	66	66	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	67	67	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	70	70	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	112	112	HYDROXYLATION (BY SIMILARITY).
FT	ACT_SITE	253	253	CHARGE RELAY SYSTEM.
FT	ACT_SITE	299	299	CHARGE RELAY SYSTEM.
FT	ACT_SITE	402	402	CHARGE RELAY SYSTEM.
FT	DISULFID	58	63	BY SIMILARITY.
FT	DISULFID	91	110	BY SIMILARITY.
FT	DISULFID	100	105	BY SIMILARITY.
FT	DISULFID	104	119	BY SIMILARITY.
FT	DISULFID	121	130	BY SIMILARITY.
FT	DISULFID	139	150	BY SIMILARITY.
FT	DISULFID	146	159	BY SIMILARITY.
FT	DISULFID	161	174	BY SIMILARITY.
FT	DISULFID	182	319	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	238	254	BY SIMILARITY.
FT	DISULFID	373	387	BY SIMILARITY.
FT	DISULFID	398	426	BY SIMILARITY.
FT	CARBOHYD	214	214	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	290	290	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	355	355	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	328	328	MISSING (IN REF. 2).
FT	CONFLICT	393	393	N -> D (IN REF. 2).
SEQ	SEQUENCE	461 AA;	51945 MW; 53FAAD0858194D6E CRC64;	

  

Alignment Scores:			
Pred. No.:	1.04e-79	Length:	461
Score:	1637.50	Matches:	289
Percent Similarity:	83.98%	Conservative:	57
Best Local Similarity:	70.15%	Mismatches:	63
Query Match:	70.07%	Indels:	3
DB:	1	Gaps:	2

  

US-09-763-153-1 (1-1245) x PRTC_MOUSE (1-461)			
QY	1	GCAACTCTCTCTCGAGGAGCTCCGTCACAGCAGCTGGAGCGGGAGTGCATAGAGGAG	60
DB	42	AlaAnSerPheLeuGluGluMetArgProGlySerLeuGluArgGluCysMetGluGlu	61
QY	61	ATCTGTGATTCGAGGAGGCCAAGGAATTTCCAAATGTGTGATGCACACTGGCCCTC	120
DB	62	IleCysAspPheGluGluAlaGluGluIlePheGlnAsnValGluAspThrLeuAlaPhe	81
QY	121	TGGTCCAGCAGCTGCAGCGTGCACCGTGTGGTCTTGCCTTGGACACCCCGTGGCC	180
DB	82	TrpIleTyrPheAspGlyAspGlnCysSerAlaProProLeuAspHisGlnCysAsp	101
QY	181	AGCCTGTGCTGGGCGACCGCAGTCATCGAGCGCATCGGCGCTTCAATTGCTCGCTG	240
DB	102	SerProCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysSerCys	141
QY	241	CGACGCGCTGGGAGGCCGCTTCTGCACGCGAGGTGAGCTTCCATTTGCTCGCTG	300
DB	122	AspLysGlyTyrGluGlyLysPheCysGlnGlnGluLeuArgPheGlnAspCysArgVal	141
QY	301	GACACGGCGCTCGACGCATTACTGCTAGAGAGGTGGGCTGCGCGCGCTGTAGCTGT	360
DB	142	AsnAsnGlyGlyCysLeuHisTyrCysLeuGluGluSerAsnGlyArgCysAlaCys	161
QY	361	GGCCTTGCTACAGCTGGGGGAGCACCTCTCTGCAGCTGCATCCCCGAGTGAAGTCCCT	420
DB	162	AlaProGlyTyrGluLeuAlaAspAspHisMetArgCysLysSerThrValAsnPhePro	181
QY	421	TGTGGAGGCCCTCGAAGCGGATGGAGAAGACGACGCTCACCTGAAACGAGACACAGAA	480
DB	182	CysGlyLysLeuGlyArgTyrPleGluGlyLysArgGlyIleLeuLysArgAspThr	200
QY	481	GACCAAGAACCAAGTA-----GATCGCGGCTCATTTGATGGGAAGATGACCAAGCGG	534

CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
CC and VIIIA.  
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D43754; BAA07811.1; -;  
DR HSSP; P04070; IPCU.  
DR MEROPS; S01.218; -;  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.  
FT NON\_TER 1 1  
FT ACT\_SITE 26 26 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 129 129 CHARGE RELAY SYSTEM.  
FT DISULFID 100 114 BY SIMILARITY.  
FT DISULFID 125 153 BY SIMILARITY.  
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT NON\_TER 161 161  
SQ SEQUENCE 161 AA; 17770 MW; 27D78F185B2FCC69 CRC64;  
  
Alignment Scores:  
Pred. No.: 2.47e-37 Length: 161  
Score: 835.00 Matches: 154  
Percent Similarity: 98.14% Conservative: 4  
Best Local Similarity: 95.65% Mismatches: 3  
Query Match: 35.73% Indels: 0  
Gaps: 1  
DB:  
  
US-09-763-153-1 (1-1245) x PRTC\_MACMU (1-161)  
  
QY 694 GAGAGTGGAGCTGACATGACATCAAGGAGGTCTTCGTCACCCCACTACAGCAAG 753  
Db 1 GluLysTrpGluLeuAspLeuAspIleGluGluValPheIleHisProAsnTrpThrLys 20  
  
QY 754 AGCACACCGACAATGACATGCTGCTGACCTGGCCCGCCGACCCCTCTCGCAG 813  
Db 21 SerThrThrAspAsnAspIleAlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerGln 40  
  
814 ACATAGTCCCATCTGCTCCCGAGACGCGCTTCGAGAGCGCGGAGCTCAATCAGGCC 873  
41 ThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluThrGlnAla 60  
  
QY 874 GCGCAGGAGACCTCGTGACGGGCTGGGCTACACAGCAGCGCAGAGAAGGAGCCCAAG 933  
Db 61 GlyGlnGluThrLeuValThrGlyTrpGlyTyrHisSerSerArgGluLysGluAlaLys 80  
  
QY 934 AGAAACCGCACCTTCTGCTCACTCACTCATCAAGATTCGCGGTCCCGCACAAATGAGTGC 993  
Db 81 ArgAsnArgThrPheIleLeuAsnPheIleLysIleProValValProArgAsnGluCys 100  
  
QY 994 ACCGAGGTATGACCAACATGTGCTGAGACATGCTGTGCGGGCATCTCGGGGAC 1053  
Db 101 SerGluValMetSerAsnMetValSerGluAsnMetLeuCysAlaGlyIleLeuGlyAsp 120  
  
QY 1054 CGCAGGATGCTGGGAGGCGACAGTGGGGGCGCCATGCTGCTCTTCACGCGCAC 1113  
Db 121 ArgGlnAspAlaCysGluGlyAspSerGlyGlyProMetValAlaSerPheHisGlyThr 140

QY 1114 TGGTTCTCTGGTGGCGCTGCTGAGCTGGGCTGAGGGCTGGGCTCTTCACAACTACGGC 1173  
Db 141 TrpPheLeuValGlyLeuValSerTrpGlyGlyGlyCysGlyLeuLeuHisAsnTrpGly 160  
  
QY 1174 GTT 1176  
Db 161 Val 161  
  
RESULT 8  
FA10\_RABIT STANDARD; PRT; 490 AA.  
ID FA10\_RABIT  
AC GI9045;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).  
GN F10.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97256311; PubMed=9101642;  
RA Pendurthi U.R., Anderson K.D., James H.L.;  
RT "Characterization of a full-length cDNA for rabbit factor X.";  
RL Thromb. Res. 85:503-514(1997).  
CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
CC converts prothrombin to thrombin in the presence of factor Va,  
CC calcium and phospholipid during blood clotting.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Thr and then  
CC Arg-I-Ile bonds in prothrombin to form thrombin.  
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
CC MORE DISULFIDE BONDS.  
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
CC CALCIUM (BY SIMILARITY).  
CC -1- PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).  
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)  
CC (BY SIMILARITY).  
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
CC ANOTHER SITE, BEYOND THE GLA DOMAIN.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF003200; AAB62542.1; -;  
DR HSSP; P00742; 1HCG.  
DR MEROPS; S01.216; -;  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002383; GLA\_blood.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR InterPro; IPR000294; VitK\_dep\_GLA.  
DR Pfam; PF00008; EGF; 2  
DR Pfam; PF00089; trypsin; 1.  
DR Pfam; PF00594; gla; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00001; GLABLOOD.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00001; EGF\_like; 1.

DR SMART: SM00069; GLA: 1.  
 DR SMART: SM00020; TRYP\_SPC: 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL: 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA: 1.  
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.  
 DR PROSITE; PS00240; TRYP\_SIN\_DOM: 1.  
 DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
 KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;  
 KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 KW Signal; Zymogen; EGF-like domain; Repeat.  
 SIGNAL 1 20  
 PROPEP 21 40  
 CHAIN 41 180  
 CHAIN 184 490  
 FT PROPEP 184 232  
 FT CHAIN 233 490  
 FT DOMAIN 86 122  
 FT DOMAIN 125 165  
 FT DOMAIN 233 490  
 FT MOD\_RES 46 46  
 FT MOD\_RES 47 47  
 FT MOD\_RES 54 54  
 FT MOD\_RES 56 56  
 FT MOD\_RES 59 59  
 FT MOD\_RES 60 60  
 FT MOD\_RES 65 65  
 FT MOD\_RES 66 66  
 FT MOD\_RES 69 69  
 FT MOD\_RES 72 72  
 FT MOD\_RES 75 75  
 FT MOD\_RES 79 79  
 FT MOD\_RES 103 103  
 FT ACT\_SITE 274 274  
 FT ACT\_SITE 320 320  
 FT ACT\_SITE 417 417  
 FT DISULFID 90 101  
 FT DISULFID 95 110  
 FT DISULFID 112 121  
 FT DISULFID 129 140  
 FT DISULFID 136 149  
 FT DISULFID 151 164  
 FT DISULFID 172 340  
 FT DISULFID 239 244  
 FT DISULFID 259 275  
 FT DISULFID 388 402  
 FT DISULFID 413 441  
 FT CARBOHYD 61 61  
 FT CARBOHYD 187 187  
 FT CARBOHYD 205 205  
 SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85AF2A6D11 CRC64;

## Alignment Scores:

Pred. No.: 2.88e-36  
 Score: 814.50  
 Percent Similarity: 53.71%  
 Best Local Similarity: 38.88%  
 Query Match: 34.85%  
 DB: 1

US-09-763-153-1 (1-1245) x FA10\_RABIT (1-490)  
 QY 1 GCCAACTCTTCCTGGAGGAGCTCGTCACAGCAGCCTGGAGCGGGAGTGCATAGAGAG 60  
 Db |||||  
 QY 41 AlaaSnSerPheLeuGluGluLeuLysGlyAsnLeuGluArgGluCysMetGluGlu 60  
 Db |||||  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCCTC 120  
 Db |||||  
 QY 61 AsnCysSerTyrGluGluAlaLeuGluValPheGluAspArgGluLysThrAsnGluPhe 80  
 Db |||||  
 QY 121 TGGTCCAAAGCAGTCGAGGAGTACCATGCTTGTGCTTGGCTTGGAGCAGCCCGTGGCC 180  
 Db |||||  
 QY 81 TrpAsnLysTyrValAspGlyAspGlnCysGluSerAsnPro 94  
 QY 181 AGCTGTGCTCGGGGACGCGCATGTCAGCGGATCGGCAGCTTCACGTCGGACTGC 240  
 Db |||||  
 QY 95 -----CysGlnAsnGlnGlyThrCysLysAspGlyLeuGlyMetTyrThrCysSerCys 112  
 Db |||||  
 QY 241 CGCAGCGGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGCTTCCTCAATTCGCTGCTG 300  
 Db |||||  
 QY 113 ValGluGlyTyrGluGlyGlnAspCysGlu---ProValThrArgLysLeuCysSerLeu 131  
 Db |||||  
 QY 301 GACACCGCGCGCTCCAGCCATTACTGCCTAGAGAGGTGGCGCGCGCTGTAGCTGT 360  
 Db |||||  
 QY 132 AspAsnGlyGlyCysAspGlnPheCysLysGluGluAsnSerValLeuCysSerCys 151  
 QY 361 GCGCTGGCTACAAAGCTGGGGGACGACCTCCTGCAGTGTCCACCCGAGTGAAGTTCCT 420  
 Db |||||  
 QY 152 AlasrGlyTyrThrLeuGlyAspAsnGlyLysSerCysIleSerThrGluLeuPhePro 171  
 QY 421 TGTGGGAGG-----CCTGGAAGCGGATGGAGAAGCGCATCAC-----462  
 Db |||||  
 QY 172 CysGlyLysValThrLeuGlyArgTyrArgArgSerProAlaThrAsnSerSerGluGly 191  
 QY 463 -----CTGAACGAGACACAGAACAGACAGAA-----489  
 Db |||||  
 QY 192 ProProGluAlaProGlyProGluGlnGlnAspGlyAsnLeuThrAlaThrGluAsn 211  
 QY 490 -----GACCAAGTAGATCG-----504  
 Db |||||  
 QY 212 ProPheAsnLeuLeuAspSerProGluProProGluAspSerSerSerLeuVal 231  
 QY 505 CGCTCATTCATGGGAAGATGACCCAGCGGGGAGACAGCCCTGGCAGGTGTCCTGCTG 564  
 Db |||||  
 QY 232 ArgIleValGlyGlyGlnAspCysArgAspGlyGluCysProTrpGlnAlaLeuVal 251  
 QY 565 GACTCAAGAAGAAGCTGGCGCTCGGGGAGTGTCTCATCCACCCCTCTGGTGTGACA 624  
 Db |||||  
 QY 252 AsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuSerGluTyrHisValLeuThr 271  
 QY 625 GCGGCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 684  
 Db |||||  
 QY 272 AlaAlaHisCysLeuHisGlnAlaLysArgPheLysValArgValGlyAspArgAspThr 291  
 QY 685 CGCGCTGGGAGAAGTGGGAGCTGGACCTGGACATCAAGAGGTCTCTCCACCCCAAC 744  
 Db |||||  
 QY 292 GluHisGluGluGlyAsnGluThrHisGluValGluValValValLysHisAsnArg 311  
 QY 745 TACAGCAAGAGCACCACCCACCAATGACATCGCATCGCTGCTGCTGCTGCTGCTGCTG 804  
 Db |||||  
 QY 312 PheValLysGluThrTyrAspPheAspIleAlaValLeuArgLeuLysThrProIleThr 331  
 QY 805 CTCTCGCAGACCATAGTGCCTATCTGCTCCGAGACAGCGGCTTTCGAGAGCGCGAGCTC 864  
 Db |||||  
 QY 332 PheArgArgAsnValAlaProAlaCysLeuProGlnLysAspTrpAlaGluSerThrLeu 351  
 QY 865 AATCAGCGCGCCAGGAGACCTCTGAGCGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 924  
 Db |||||  
 QY 352 ---MetAlaGlnLysThrGlyIleValSerGlyPheGly-----ArgThrHis 366  
 QY 925 GAGGCGCAAGAGAACCGGACCTTCTGCTCAACTTCATCAAGATTCCTGCTCCGCGCAC 984  
 Db |||||





CC MORE DISULFIDE BONDS.  
 CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -!- PTM: N- AND O-GLYCOSYLATED.  
 CC -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X00673; CAA25286.1; .  
 CC PIR: A00925; EXBO.  
 CC PDB: IAP0; 31-JAN-94.  
 CC PDB: ICCF; 31-MAY-94.  
 CC PDB: IWHE; 15-MAY-97.  
 CC PDB: IWHF; 15-MAY-97.  
 CC MEROPS: S01.216; .  
 CC GlycoSuiteDB: P00743; .  
 CC InterPro: IPR000152; Asx\_hydroxyl.  
 CC InterPro: IPR0001314; Chymotrypsin.  
 CC InterPro: IPR0000561; EGF-like.  
 CC InterPro: IPR000742; EGF-2.  
 CC InterPro: IPR001881; EGF\_Ca.  
 CC InterPro: IPR002383; GLA\_blood.  
 CC InterPro: IPR001254; Ser\_protease\_Try.  
 CC InterPro: IPR000294; VitK\_dep\_GLA.  
 CC Pfam: PF00089; trypsin; 1.  
 CC Pfam: PF00594; gla; 1.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC PRINTS: PR00001; GLABLOOD.  
 CC SMART: SM00179; EGF\_CA; 1.  
 CC SMART: SM00001; EGF\_like; 1.  
 CC SMART: SM00069; GLA; 1.  
 CC SMART: SM00020; TRYD\_SPC; 1.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 CC PROSITE: PS00022; EGF\_1; 1.  
 CC PROSITE: PS01187; EGF\_CA; 1.  
 CC PROSITE: PS01186; EGF\_2; 2.  
 CC PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE: PS50240; TRYPSIN\_DOM; 1.  
 CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC Glycoprotein: Hydrolase; Serine protease; Plasma; Blood coagulation;  
 CC Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 CC Signal; Zymogen; EGF-like domain; Repeat; Sulfation; 3D-structure.  
 CC SIGNAL 1 23 POTENTIAL.  
 CC PROPEP 24 40  
 CC CHAIN 41 180 FACTOR X LIGHT CHAIN.  
 CC CHAIN 183 492 FACTOR X HEAVY CHAIN.  
 CC PROPEP 183 233 ACTIVATION PEPTIDE.  
 CC CHAIN 234 492 ACTIVATED FACTOR XA.  
 CC PROPEP 476 492 MAY BE REMOVED BUT IS NOT NECESSARY FOR  
 CC ACTIVATION.  
 CC DOMAIN 86 122 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 125 165 EGF-LIKE 2.  
 CC DOMAIN 234 492 SERINE PROTEASE.  
 CC ACT\_SITE 275 275 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 321 321 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 418 418 CHARGE RELAY SYSTEM.  
 CC MOD\_RES 46 46 GAMMA-CARBOXYGLUTAMIC ACID.  
 CC MOD\_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID.  
 CC MOD\_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID.

FT	MOD_RES	56	56	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	59	59	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	66	66	GAMMA-CARBOXYGLUTAMIC ACID.

Alignment Scores: 5.62e-36 Length: 492  
 Pred. No.: 809.00 Matches: 171  
 Score: 52.27% Conservative: 71  
 Percent Similarity: 36.93% Mismatches: 149  
 Best Local Similarity: 34.62% Indels: 72  
 Query Match: 1 Gaps: 12  
 DB: 1

US-09-763-153-1 (1-1245) x FA10\_BOVIN (1-492)

QY	1	GCCAACTCTCTCTGGAGGAGCTCCGTCACAGCAGCGCTGGAGGGAGTGTCATAGAGGAG	60
Db	41	AlaAsnSerPheLeuGluGluValLysGlnGlyAsnLeuGluArgGluCysLeuGluGlu	60
QY	61	ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC	120
Db	61	AlaCysSerLeuGluGluAlaArgGluValPheGluAlaGluGlnThrAspGluPhe	80
QY	121	TGGTCCAAAGCAGCTCGACGGTGCAGCTGCTTGGTCTTGGCTTGGAGCACCCTGGCC	180
Db	81	TrpSerLysTyrLysAspGlyAspGlnCys	96
QY	181	AGCCTGTGCTGCGGGCAGCGCAGCTGCATCGACGGCATCGCAGCTTCTAGCTCGACTGC	240
Db	97	Asn	112
QY	241	CGCAGCGGCTGGAGGGCCCTCTTCCAG	291
Db	113	AlaGluGlyPheGluGlyLysAsnCysGluPheSerThrArgGluLeu	128
QY	292	TGCTCGCTGGACAACGGCGCTGCACGCATTACTGCTAGAGGAGTGGCTGGCGGCGC	351
Db	129	CysSerLeuAspAsnGlyGlyCysAspGlnPheCysArgGluGluArgSerGluValArg	148
QY	352	TGTAGCTGTGCGCTGGCTACAAAGCTGGGGGACAGCTCTCTGCAGTGTCCACCCGCGAGTG	411
Db	149	CysSerCysAlaHisGlyTyrValLeuGlyAspSerLysSerCysValSerThrGlu	168
QY	412	AAGTTCCTTGTGGAGG---CCCTGGAAGCGGATGAGAAGAGCGCATCTACCTGAAA	468
Db	169	ArgPheProCysGlyLysPheThrGlnGlyArgSerArgArgTrpAlaIleHisThrSer	188
QY	469	CGAGACACAGAACAC-----CAAGAGACCAAGTAGATCCG-----	504
Db	189	GluAspAlaLeuAspAlaSerGluLeuGluHisTyrAspProAlaAspLeuSerProThr	208
QY	504	-----	504
Db	209	GluSerSerLeuAspLeuLeuGlyLeuAsnArgThrGluProSerAlaGlyGluAspGly	228
QY	505	-----CGGCTCATTTGATGGAGAGATGACCGGGGGGAGACAGCCCTGGGAG	552
Db	229	SerGlnValValArgIleValGlyArgAspCysAlaGluGlyGluCysProThrGln	248
QY	553	GTGGTCTCTGCTGGACTCAAGAGAGAGCTGGCTGGGGGGCAGTGTCTATCCACCCCTCC	612
Db	249	AlaLeuLeuValAsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuAsnGluPhe	268
QY	613	TGGGTCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	672
Db	269	TyrValLeuThrAlaAlaHisCysLeuHisGlnAlaLysArgPheThrValArgValGly	288
QY	673	GAGTATGACCTGGCGGCTGGGAGAGAGTGGGAGCTGGACATCAAGAGAGGTCTTC	732
Db	289	AspArgAsnThrGluGlnGluGluGlyAsnGluMetAlaHisGluValGluMetThrVal	308
QY	733	GTCCACCCCACTACAGACAGACCCACCGACATGACATGCTGCTGCTGCTGCTGCTGCTG	792



217 DISULFID 233 BY SIMILARITY.  
 349 DISULFID 368 BY SIMILARITY.  
 379 DISULFID 407 BY SIMILARITY.  
 45 MOD\_RES 45 GAMMA-CARBOXYGLUTAMIC ACID.  
 46 MOD\_RES 46 GAMMA-CARBOXYGLUTAMIC ACID.  
 53 MOD\_RES 53 GAMMA-CARBOXYGLUTAMIC ACID.  
 55 MOD\_RES 55 GAMMA-CARBOXYGLUTAMIC ACID.  
 58 MOD\_RES 58 GAMMA-CARBOXYGLUTAMIC ACID.  
 59 MOD\_RES 59 GAMMA-CARBOXYGLUTAMIC ACID.  
 64 MOD\_RES 64 GAMMA-CARBOXYGLUTAMIC ACID.  
 65 MOD\_RES 65 GAMMA-CARBOXYGLUTAMIC ACID.  
 68 MOD\_RES 68 GAMMA-CARBOXYGLUTAMIC ACID.  
 74 MOD\_RES 74 GAMMA-CARBOXYGLUTAMIC ACID.  
 102 MOD\_RES 102 HYDROXYLATION (BY SIMILARITY).  
 211 CARBOHYD 211 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 242 CARBOHYD 242 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 306 CARBOHYD 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 444 AA; 49011 MW; 0481ABC4FE5427F8 CRC64;  
 SEQUENCE

Alignment Scores:  
 Pred. No.: 1.32e-35 Length: 444  
 Score: 802.00 Matches: 170  
 Percent Similarity: 56.03% Conservative: 67  
 Best Local Similarity: 40.19% Mismatches: 144  
 Query Match: 34.32% Indels: 42  
 Gaps: 11  
 DB: 1

US-09-763-153-1 (1-1245) x FA7\_RABIT (1-444)  
 1 GCCAACTCCTCTCTCGAGAGAGCTCGCTCACAGACGCTGCGAGCGGAGTGCATAGAGAG 60  
 40 AlaAsnSerPheLeuGluGluLeuArgProGlySerLeuGluArgGlyCysLeuGlu 59  
 61 ATCTGTGACTTCCAGAGCGCAAGAAATTTTCCAAAATGTGGATGACACTGGCTTC 120  
 60 LeuCySerPheGluGluAlaArgGluValPheGlnSerThrGluArgThrLysGlnPhe 79  
 121 TGGTCCAAAGCAGCTCGACGGTGACAGTCTTGGCTTGGCTTGGAGCACCCGTGGGCC 180  
 80 TrpIleThrTyrAsnAspGlyAspGlnCysAlaSerAsnPro----- 93  
 181 AGCTGTGCTCGGGCAGCGACGCTGCATCGACGGGATCGGAGCTTCAGTTCGACTGC 240  
 94 -----CysGlnAsnGlyGlySerCysGluAspGlnIleGlnSerTyrlleCysPheCys 111  
 241 CGCAGCGCTGGAGGCGCGCTCTCGACGGGAGGTGAGC---TTCTCAATTTGCTCG 297  
 112 LeuAlaAspPheGluGlyArgAsnCyGluGlyAsnLysAsnAspGlnLeuIleCysMet 131  
 298 CTGGACAAGCGGCTGCACCATCTACTGCTAGAGAGGTGGCTGGCGGC---TGT 354  
 132 TyrGluAsnGlyGlyCysGluGlnTyrCysSerAspHisValGlySerGlnArgSerCys 151  
 355 AGCTGTGCGCCTGTACAGTGGGGAGCACTCTCGCAGTGTCCACCCCGCAGTGAG 414  
 152 ArgCysHisGluGlyTyrThrLeuLeuProAsnGlyValSerCysThrProThrValasp 171  
 415 TTCCCTTGTGGAGGCGCTTGAAGCGGATGGAGAGAGCGCCAGTCCACTGAAACGAGAC 474  
 172 TyrProCysGlyLysVal---ProAlaLeuGluLysArgGlyAla----- 185  
 475 ACAGAAGACCAAGAGACCAAGTAGATCGCGGCTCATTTGATGGGAAGATGACACGCGG 534  
 186 -----SerAsnProGlnGlyArgIleValGlyGlyLysValCysProLys 200  
 535 GGAGACAGCCCTGGCAGGTGGTCTGCTGGACTCAAGACAGACGCTGGCTGGCGGCA 594  
 201 GlyGluCysProTrpGlnAlaAlaLeuMetAsnGly---SerThrLeuLeuCysGlyGly 219  
 595 GTGCTCATCCACCCCTCTCTGGTCTGCACAGCGGCGCCACTGCATGATGAG- 645  
 645 GTGCTCATCCACCCCTCTCTGGTCTGCACAGCGGCGCCACTGCATGATGAG- 645

646 TCCAAAGAGCTCCTTGTACAGCTTGGAGAGTATGACCTGGCGCTGGGAGAGTGGGAG 705  
 240 LeuArgAsnLeuThrIleValLeuGlyGluHisAspLeuSerGluHisGluGlyAspGlu 259  
 706 CTGGACCTGGACATCAAGGAGGTCTTCGTCCTCCACCCCACTACAGCAGAGACCCACCGAC 765  
 260 GlnValArgHisAlaGlnLeuIleMetProAspLysTyrValProGlyLysThrAsp 279  
 766 AATGACATCGCACTCTGTCACCTGGCCCGCCAGCCCACTCTCGACAGCACCATTAGTGCC 825  
 280 HisAspIleAlaLeuLeuArgLeuGlnProAlaAlaLeuThrAsnAsnValValPro 299  
 826 ATCTGCTCCCGAGCAGCGCTTCGACAGCGCGAGCTCAATCAGCCCGCCAGGAGACC 885  
 300 LeuCysLeuProGluArgAsnPheSerGluSerThrLeuAlaThrIle---ArgPheSer 318  
 886 CTGCTGACGCGGTGGGGC-----TACCACAGCAGCGAGAGAGAGGAGGAGCAGAGA 936  
 319 ArgValSerGlyTrpGlyGlnLeuLeuTyrArgGlyAlaLeuAlaArgGlu----- 335  
 937 AACGCGACCTTCCTCAACTTCATCAAGATTCCCGTGGTCCGCGACATGAGTGCAGC 996  
 336 -----LeuMetAlaIleAspValProArgLeuMetThrGlnAspCysVal 350  
 997 GAGGTGCTG-----ACCACATGGTGTCTGAGAACATGCTGTGTGGCGGC 1041  
 351 GluGlnSerGluHisLysProGlySerProGluValThrGlyAsnMetPheCysAlaGly 370  
 1042 ATCTCTGGGAGCCGCGAGGATGCTCGAGGGCGCAGTGGGGGCCCATGTCGCTCC 1101  
 371 TyrLeuAspGlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrSer 390  
 1102 TTCCACGGCACCTGTTCTCTGGTGGGCTGTGAGCTGGGTGGAGGTGGGTGGCTT 1161  
 391 TyrHisGlyThrTyrThrLeuThrGlyValValSerTrpGlyGlyCysAlaVal 410  
 1162 CACAACCTAGCGGTTCACACAAAGTACAGCGCTTACCTCGACTGGATTCATGGCACATC 1221  
 411 GlyHisValGlyValTyrThrArgValSerArgTyrThrGluTrpLeuSerArgLeuMet 430  
 1222 AGAGACAAG 1230  
 431 ArgSerLys 433

RESULT 11  
 FA10\_CHICK STANDARD; PRT: 475 AA.  
 ID FA10\_CHICK  
 AC P25155;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor)  
 DE (Virus activating protease) (VAP).  
 GN FX.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Chorioallantoic membrane;  
 RX MEDLINE=91257322; PubMed=2044767;  
 RA Suzuki H., Harada A., Hayashi Y., Wada K., Asaka J.-I., Gotoh B.,  
 RA Ogasawara T., Nagai Y.;  
 RA "Primary structure of the virus activating protease from chick  
 RA embryo. Its identity with the blood clotting factor Xa.";  
 RT FEBS Lett. 283:281-285(1991).  
 RL [2]  
 RN SEQUENCE OF 41-55 AND 241-261.  
 RP TISSUE=Allantoic fluid;  
 RC MEDLINE=91065352; PubMed=2174359;  
 RX Gotoh B., Ogasawara T., Toyoda T., Innocencio N.M., Hamaguchi M.,  
 RA

RA Negai Y.;  
 RT "An endoprotease homologous to the blood clotting factor X as a  
 RL EMBO J. 9:4189-4195(1990).  
 CC -|- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
 CC converts prothrombin to thrombin in the presence of factor Va,  
 CC calcium and phospholipid during blood clotting.  
 CC -|- FUNCTION: VAP CLEAVES THE FUSION PROTEINS OF SENDAI VIRUS, NDV,  
 CC AND INFLUENZA VIRUS A AT A SPECIFIC SINGLE ARGinine-CONTAINING  
 CC SITE, AND PLAYS A KEY ROLE IN THE VIRAL SPREADING IN THE ALLANTOIC  
 CC SAC.  
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then  
 CC Arg-|-Ile bonds in prothrombin to form thrombin.  
 CC -|- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
 CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
 CC MORE DISULFIDE BONDS.  
 CC -|- TISSUE SPECIFICITY: LIVER AND CHORIOALLANTOIC MEMBRANE.  
 CC -|- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -|- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -|- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: D00844; BAA00724.1; -  
 DR PIR: S15838; S15838.  
 DR HSP: P00742; IHCG.  
 DR MEROPS: S01.216; -  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001234; Ser\_protease\_Try.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00010; EGF\_BLOOD.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_CA; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 DR Glycoprotein: Hydrolase; Serine protease; Plasma; Blood coagulation;  
 KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 KW Signal; Zymogen; EGF-like domain; Repeat.  
 FT SIGNAL 1 20 OR 30, OR 31 (POTENTIAL).  
 FT PROPEP 21 40  
 FT CHAIN 41 180  
 FT CHAIN 186 475  
 FT PROPEP 186 475  
 FT CHAIN 242 475  
 FT CHAIN 86 122  
 FT DOMAIN

FT DOMAIN	125	168	EGF-LIKE 2.
FT DOMAIN	241	475	SERINE PROTEASE
FT MOD_RES	46	46	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	47	47	(BY SIMILARITY)
FT MOD_RES	47	47	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	54	54	(BY SIMILARITY)
FT MOD_RES	54	54	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	56	56	(BY SIMILARITY)
FT MOD_RES	56	56	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	59	59	(BY SIMILARITY)
FT MOD_RES	59	59	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	60	60	(BY SIMILARITY)
FT MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	65	65	(BY SIMILARITY)
FT MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	66	66	(BY SIMILARITY)
FT MOD_RES	66	66	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	69	69	(BY SIMILARITY)
FT MOD_RES	69	69	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	72	72	(BY SIMILARITY)
FT MOD_RES	72	72	GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES	79	79	(BY SIMILARITY)
FT MOD_RES	79	79	GAMMA-CARBOXYGLUTAMIC ACID
FT ACT_SITE	103	103	HYDROXYLATION (BY SIMILARITY).
FT ACT_SITE	282	282	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE	328	328	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE	425	425	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID	90	101	BY SIMILARITY.
FT DISULFID	95	110	BY SIMILARITY.
FT DISULFID	112	121	BY SIMILARITY.
FT DISULFID	129	140	BY SIMILARITY.
FT DISULFID	136	152	BY SIMILARITY.
FT DISULFID	154	167	BY SIMILARITY.
FT DISULFID	175	348	INTERCHAIN (BY SIMILARITY).
FT DISULFID	247	252	BY SIMILARITY.
FT DISULFID	267	283	BY SIMILARITY.
FT DISULFID	396	410	BY SIMILARITY.
FT CARBOHYD	421	449	BY SIMILARITY.
FT CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	207	207	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	228	228	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	285	285	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE	475 AA;	53142 MW;	570BF84956C5C74D CRG64;

Alignment Scores:  
 Pred. No.: 1.4e-35  
 Score: 801.50  
 Percent Similarity: 54.77%  
 Best Local Similarity: 36.14%  
 Query Match: 34.30%  
 DB: 1  
 Gaps: 8

US-09-763-153-1 (1-1245) x FA10\_CHICK (1-475)

QY 1	CCCAACTCTCTCTGGAGGAGCTCCGTCACACAGCTGGAGCGGAGTGATGATAGGAG 60
Db 41	AlaAsnSerPheLeuGluGluMetLysGlnGlyAsnIleGluArgGluCysAsnGluGlu 60
QY 61	ATCTGTGACTTCGAGGAGCCCAAGGAATTTCCAAATGTGGATGACACTGGCCTTC 120
Db 61	ArgCysSerLysGluGluAlaArgGluAlaPheGluAspAsnGluLysThrGluGluPhe 80
QY 121	TGTTCACAGCTCGACGGTGACAGTGCTGTGCTGCTGCTGGAGCACCCTCGCC 180
Db 81	TrpAsnIleTyrValAspGlyAspGlnCysSerSerAsnPro----- 94
QY 181	AGCCTGTGTCGGCGGACGACGCTGCTGATCAGCGATCGGACGCTGACGCTGCG 240
Db 95	-----CysHisTyrGlyGlyGlnCysLysAspGlySerTyrThrCysSerCys 112
QY 241	CCGACGCGCTGGAGGCGCGCTTCTGCCACGCGAGGTGAGCTTCTCAATTCGCTG 300

Db 113 LeuAspGlyTyrGluGlyLysAsnCysGluPheValIlePro---LysTyrCysLysIle 131  
QY 301 GACACGGCGCTGCACGCATTTACTGCTAGAGAGAGTGGCTGGCGGC----- 351  
Db 132 AsnAsnGlyAspCysGluGlnPheCysSerIleLysLysSerValGlnLysAspValVal 151  
QY 352 TGTAGCTGTGGCTGCTACAGCTGGGGAGACCTCTCTGAGTGTACCCCGGAGTG 411  
Db 152 CysSerCysThrSerGlyTyrGluLeuAlaGluAspGlyLysGlnCysValSerLysVal 171  
QY 412 AAGTTCCTCTGTGGAGCCCTGGAGCGGATGGAGAGAGCGAGTCACTCAAAACGA 471  
Db 172 LysTyrProCysGlyLysValLeuMetLysArgIleLysArgSerValIleLeuProThr 191  
QY 472 GAC-----ACAGAGACCAA----- 486  
Db 192 AsnSerAsnThrAsnAlaThrSerAspGlnAspValProSerThrAsnGlySerIleLeu 211  
QY 486 ----- 486  
Db 212 GluGluValPheThrThrThrThrGluSerProThrProProArgAsnGlySerSer 231  
QY 487 -----GAAGACCAAGTAGATCCCGCTCATTCATGGGAAGATGACCGCGGGAGAC 540  
Db 232 IleThrAspProAsnValAsnThrArgIleValGlyGlyAspGluCysArgProGlyGlu 251  
QY 541 AGCCCTGGCAGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 252 CysProThrGlnAlaValLeuIleAsnGluLysGlyGluGluPheCysGlyGlyThrIle 271  
QY 601 ATCCACCCCTCTCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 272 LeuAsnGluAspPheIleLeuThrAlaAlaHisCysIleAsnGlnSerLysGluIleLys 291  
QY 661 GTCAGCTTGAGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 292 ValValValGluValValAspArgGluLysGluGluHisSerGluThrHisThrAla 311  
QY 721 AAGAGGTCTGCTGCCACCCCACTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 312 GluLysIlePheValHisSerLysTyrIleAlaGluThrTyrAspAsnAspIleAlaLeu 331  
QY 781 CTGACCTGCGCCAGCGCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 332 IleLysLeuLysGluProIleGlnPheSerGluThrValValProAlaCysLeuProGln 351  
QY 841 AGCGCCCTTGACAGCGCGAGCTC---AATCAGCGCCGCCAGAGAGAGAGAGAGAGAGAGAG 897  
Db 352 AlaAspPheAlaAsnGluValLeuMetAsnGln-----LysSerGlyMetValSerGly 369  
QY 898 TGGGCTACCCAGCAGCGGAG 957  
Db 370 PheGlyArgGluPheGluAlaGlyArgLeuSerLysArg-----LeuLys 384  
QY 958 TTTCACAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017  
Db 385 ValLeuGluValProTyrValAspArgSerThrCysLysGlnSerThrAsnPheAlaIle 404  
QY 1018 TCTGAGACATGCTGTGTGGGGGATCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077  
Db 405 ThrGluAsnMetPheCysAlaGlyTyrGluThrGluGlnLysAspAlaCysGlnGlyAsp 424  
QY 1078 AGTGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
Db 425 SerGlyGlyProHisValThrArgTyrLysAspThrThrPheValThrGlyIleValSer 444  
QY 1138 TGGGGTGGAGGCTGTGGGCTCTCTACAACTACGGCGTTTACAAAGTACAGCGCTGCTGCTGCTGCT 1197  
Db 445 TrpGlyGluGlyCysAlaArgLysGlyLysTyrGlyValTyrThrLysLeuSerArgPhe 464  
QY 1198 CTCAGCTGATCATGGGACATCATGAGACAG 1230  
Db 465 LeuArgTrpValArgThrValMetArgGlnLys 475

RESULT 12  
FA10\_HUMAN  
ID FA10\_HUMAN STANDARD; PRT; 488 AA.  
AC P00742; Q14340;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).  
GN F10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91216473; PubMed=1902434;  
RA Messler T.L., Pittman D.D., Long G.L., Kaufman R.J., Church W.R.;  
RT "Cloning and expression in COS-1 cells of a full-length cDNA encoding  
human coagulation factor X";  
RL Gene 99:291-294(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87026600; PubMed=3768336;  
RA Leytus S.P., Foster D.C., Kurachi K., Davie E.W.;  
RT "Gene for human factor X: a blood coagulation factor whose gene  
organization is essentially identical with that of factor IX and  
protein C";  
RL Biochemistry 25:5098-5102(1986).  
RN [3]  
RP SEQUENCE OF 13-488 FROM N.A.  
RX MEDLINE=85216545; PubMed=2582420;  
RA Fung M.R., Hay C.W., McGillivray R.T.A.;  
RT "Characterization of an almost full-length cDNA coding for human  
blood coagulation factor X";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:3591-3595(1985).  
RN [4]  
RP SEQUENCE OF 19-488 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=86221713; PubMed=3011603;  
RA Kaul R.K., Hildebrand B., Roberts S., Jagadeeswaran P.;  
RT "Isolation and characterization of human blood-coagulation factor X  
cDNA";  
RL Gene 41:311-314(1986).  
RN [5]  
RP SEQUENCE OF 41-179  
RX MEDLINE=83257207; PubMed=6871167;  
RA McMullen B.A., Fujikawa K., Kisiel W., Sasagawa T., Howald W.N.,  
Kwa E.Y., Weinstein B.;  
RT "Complete amino acid sequence of the light chain of human blood  
coagulation factor X: evidence for identification of residue 63 as  
beta-hydroxyaspartic acid";  
RL Biochemistry 22:2875-2884(1983).  
RN [6]  
RP SEQUENCE OF 115-488 FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=Liver;  
RX MEDLINE=84222026; PubMed=6587384;  
RA Leytus S.P., Chung D.W., Kisiel W., Kurachi K., Davie E.W.;  
RT "Characterization of a cDNA coding for human factor X";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:3699-3702(1984).  
RN [7]  
RP SEQUENCE OF 183-234, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=94062825; PubMed=8243461;  
RA Inoue K., Morita T.;  
RT "Identification of O-linked oligosaccharide chains in the activation  
peptides of blood coagulation factor X. The role of the carbohydrate  
moieties in the activation of factor X";  
RL Eur. J. Biochem. 218:153-163(1993).  
RN [8]  
RP SEQUENCE OF 1-23 FROM N.A.  
RX MEDLINE=90128299; PubMed=2612918;  
RA Jagadeeswaran P., Reddy S.V., Rao K.J., Hamsabhushanam K., Lyman G.;  
RT "Cloning and characterization of the 5' end (exon 1) of the gene

encoding human factor X.  
 Gene 84:517-519(1989).  
 [9]  
 X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-179 AND 235-278.  
 MEDLINE-93360277; PubMed-8355279;  
 Padmanabhan K., Padmanabhan K.P., Tulinsky A., Park C.H., Bode W.,  
 Huber R., Blankenship D.T., Cardin A.D., Kistler W.,  
 "Structure of human Des(1-45) factor Xa at 2.2-A resolution."  
 J. Mol. Biol. 232:947-966(1993).  
 [10]  
 X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-179 AND 235-278.  
 MEDLINE-98283982; PubMed-9618463;  
 Kamata K., Kawamoto H., Homma T., Iwama T., Kim S.H.;  
 "Structural basis for chemical inhibition of human blood coagulation  
 factor Xa."  
 Proc. Natl. Acad. Sci. U.S.A. 95:6630-6635(1998).  
 -|- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
 converts prothrombin to thrombin in the presence of factor Va,  
 calcium and phospholipid during blood clotting.  
 -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then  
 Arg-|-Ile bonds in prothrombin to form thrombin.  
 -|- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
 BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
 MORE DISULFIDE BONDS.  
 -|- TISSUE SPECIFICITY: Plasma; synthesized in the liver.  
 -|- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME  
 GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CALCIUM.  
 -|- PTM: N- AND O-GLYCOSYLATED.  
 -|- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 -|- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; K03194; AAA52490.1; -;  
 EMBL; M57285; AAA52421.1; -;  
 EMBL; L29433; AAA52764.1; -;  
 EMBL; L00390; AAA52764.1; JOINED.  
 EMBL; L00391; AAA52764.1; JOINED.  
 EMBL; L00392; AAA52764.1; JOINED.  
 EMBL; L00393; AAA52764.1; JOINED.  
 EMBL; L00394; AAA52764.1; JOINED.  
 EMBL; L00395; AAA52764.1; JOINED.  
 EMBL; L00396; AAA52764.1; JOINED.  
 EMBL; M22613; AAA51984.1; -;  
 EMBL; K01886; AAA52486.1; -;  
 EMBL; M33297; AAA52636.1; -;  
 PIR; A00924; EXHU.  
 PIR; A25853; A25853.  
 PIR; A24478; A24478.  
 PDB; IHCG; 08-MAY-95.  
 PDB; IFAX; 29-OCT-97.  
 PDB; IFXY; 17-JUN-98.  
 PDB; IXKA; 23-MAR-99.  
 PDB; IXKB; 23-MAR-99.  
 MEROPS; S01.216; -;  
 GlycoSuiteDB; P00742; -;  
 Genew; HGNC:3528; F10.  
 MIM; 134530; -;  
 MIM; 227600; -;  
 InterPro; IPR000152; Asx\_hydroxyl.  
 InterPro; IPR001314; Chymotrypsin.  
 InterPro; IPR000561; EGF-like.  
 InterPro; IPR000742; EGF\_2.  
 InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; Vitk\_dep\_GLA.  
 DR Pfam; PF00008; EGF\_2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00594; gla; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF\_like; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00101; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;  
 KW gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 KW Signal; zymogen; EGF-like domain; Repeat; 3D-structure.  
 FT SIGNAL 1 31  
 FT PROPEP 32 40  
 FT CHAIN 41 179  
 FT CHAIN 183 488  
 FT PROPEP 183 234  
 FT CHAIN 235 488  
 FT DOMAIN 86 122  
 FT DOMAIN 125 165  
 FT DOMAIN 235 488  
 FT MOD\_RES 46 46  
 FT MOD\_RES 47 47  
 FT MOD\_RES 54 54  
 FT MOD\_RES 56 56  
 FT MOD\_RES 59 59  
 FT MOD\_RES 60 60  
 FT MOD\_RES 65 65  
 FT MOD\_RES 66 66  
 FT MOD\_RES 69 69  
 FT MOD\_RES 72 72  
 FT MOD\_RES 79 79  
 FT MOD\_RES 103 103  
 FT CARBOHYD 199 199  
 FT CARBOHYD 211 211  
 FT CARBOHYD 221 221  
 FT CARBOHYD 231 231  
 FT ACT\_SITE 276 276  
 FT ACT\_SITE 322 322  
 FT ACT\_SITE 419 419  
 FT DISULFID 90 101  
 FT DISULFID 95 110  
 FT DISULFID 112 121  
 FT DISULFID 129 140  
 Alignment Scores:  
 Pred. No.: 14e-35 Length: 488  
 Score: 801.50 Matches: 159  
 Percent Similarity: 54.55% Conservative: 87  
 Best Local Similarity: 35.25% Mismatches: 150  
 Query Match: 34.30% Indels: 55  
 DB: 1 Gaps: 8

US-09-763-153-1 (1-1245) x FA10\_HUMAN (1-488)

QY 1 GCCAACTCCTTCTCGAGGAGCTCCGTACAGCAGCGGAGGAGTCATAGAGGAG 60  
 |||||||||||||||||||  
 Db 41 AlaAsnSerPheLeuGluGluMetLysGlyHisLeuGluArgLysCysMetGluGlu 60  
 |||||||||||||||||||  
 QY 61 ATCTGTGACTTCGAGGAGGCCCAAGGAATTTCCAAAATGTGGATGACACACTGGCCTTC 120



[illegible]

RT (Xyl2-Glc) O-glycosidically linked to a serine residue in the first  
 RT epidermal growth factor-like domain of human factors VII and IX and  
 RL protein Z and bovine protein Z.";  
 RN J. Biol. Chem. 264:20320-20325(1989).  
 RN [7]  
 RP STRUCTURE OF CARBOHYDRATE ON SER-112.  
 RX MEDLINE=91344709; PubMed=2129367;  
 RA Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;  
 RA "A new trisaccharide sugar chain linked to a serine residue in the  
 RT first EGF-like domain of clotting factors VII and IX and protein Z.";  
 RN Adv. Exp. Med. Biol. 281:121-131(1990).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.  
 RX MEDLINE=96175641; PubMed=8598903;  
 RA Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,  
 RA Konigsberg W.H., Nemerson Y., Kirchhofer D.;  
 RA "The crystal structure of the complex of blood coagulation factor  
 RT VIIa with soluble tissue factor.";  
 RN Nature 380:41-46(1996).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.  
 RX MEDLINE=99126538; PubMed=9925787;  
 RA Zhang E., St Charles R., Tulinsky A.;  
 RA "Structure of extracellular tissue factor complexed with factor VIIa  
 RT inhibited with a BPTI mutant.";  
 RN J. Mol. Biol. 285:2089-2104(1999).  
 RN [10]  
 RP STRUCTURE BY NMR OF 105-145.  
 RX MEDLINE=98367502; PubMed=9692950;  
 RA Muranyi A., Finn B.E., Gippert G.P., Forsen S., Stenflo J.,  
 RA Drakenberg T.;  
 RA "Solution structure of the N-terminal EGF-like domain from human  
 RT factor VII.";  
 RN Biochemistry 37:10605-10615(1998).  
 RN [11]  
 RP VARIANTS GLN-364.  
 RX MEDLINE=91300046; PubMed=2070047;  
 RA O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,  
 RA Meade T.W., Tuddenham E.G.D.;  
 RA "Purification and characterization of factor VII 304-Gln: a variant  
 RT molecule with reduced activity isolated from a clinically unaffected  
 RN male.";  
 RL Blood 78:132-140(1991).  
 RN [12]  
 RP VARIANTS GLN-364 AND PHE-370.  
 RX MEDLINE=92340074; PubMed=1634227;  
 RA Marchetti G., Patraccini P., Gemmati D., Derosa V., Pinotti M.,  
 RA Rodorigo G., Casonato A., Girolami A., Bernardi F.;  
 RA "Detection of two missense mutations and characterization of a repeat  
 RT polymorphism in the factor VII gene (F7).";  
 RL Hum. Genet. 89:497-502(1992).  
 RN [13]  
 RP VARIANT TYR-238.  
 RX MEDLINE=93372811; PubMed=8364544;  
 RA Marchetti G., Ferrati M., Patraccini P., Redaelli P., Bernardi F.;  
 RA "A missense mutation (178Cys-->Tyr) and two neutral dimorphisms  
 RT (115His and 333Ser) in the human coagulation factor VII gene.";  
 RN Hum. Mol. Genet. 2:1055-1056(1993).  
 RN [14]  
 RP VARIANTS.  
 RX MEDLINE=94061028; PubMed=8242057;  
 RA Takamiya O., Kemball-Cook G., Marin D.M.A., Cooper D.N.,  
 RA von Felten A., Meili B., Hahn I., Prangnell D.R., Lumley H.,  
 RA Tuddenham E.G.D., McVey J.H.;  
 RA "Detection of missense mutations by single-strand conformational  
 RT polymorphism (SSCP) analysis in five dysfunctional variants of  
 RL coagulation factor VII.";  
 RN Hum. Mol. Genet. 2:1355-1359(1993).  
 RN [15]  
 RP VARIANTS CHARLOTTE GLN-139 AND GLN-212.  
 RX MEDLINE=94264305; PubMed=8204879;  
 RA Chaign S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,  
 RA Roberts H.R., Blajchman M., Monroe D.M., High K.A.;

RT "Severe factor VII deficiency caused by mutations abolishing the  
 RT cleavage site for activation and altering binding to tissue factor.";  
 RL Blood 83:3524-3535(1994).  
 RN [16]  
 RP VARIANT VAL-354.  
 RX MEDLINE=95072589; PubMed=7981691;  
 RA Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,  
 RA Rodeghiero F., Marchetti G.;  
 RA "Topologically equivalent mutations causing dysfunctional coagulation  
 RT factors VII (294Ala-->Val) and X (334Ser-->Pro).";  
 RN Hum. Mol. Genet. 3:1175-1177(1994).  
 RN [17]  
 RP VARIANT MIE HIS-307.  
 RX MEDLINE=95064662; PubMed=7974346;  
 RA Ohiwa M., Hayashi T., Wada H., Minamikawa K., Shirakawa S.,  
 RA Suzuki K.;  
 RA "Factor VII MIE: homozygous asymptomatic type I deficiency caused by  
 RT an amino acid substitution of His (CAC) for Arg(247) (CGC) in the  
 RL catalytic domain.";  
 RN Thromb. Haemost. 71:773-777(1994).  
 RN [18]  
 RP VARIANT MET-419.  
 RX MEDLINE=96247510; PubMed=8652821;  
 RA Arbini A.A., Mannucci P.M., Bauer K.A.;  
 RA "A Thr359Met mutation in factor VII of a patient with a hereditary  
 RT deficiency causes defective secretion of the molecule.";  
 RN Blood 87:5085-5094(1996).  
 RN [19]  
 RP VARIANTS W-283; K-325; V-358; Q-364; E-402 AND Q-413.  
 RX MEDLINE=97001216; PubMed=8844208;  
 RA Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,  
 RA Lunghi B., Rodeghiero F., Marchetti G.;  
 RA "Mutation pattern in clinically asymptomatic coagulation factor VII  
 RT deficiency.";  
 RL Hum. Mutat. 8:108-115(1996).  
 RN [20]  
 RP VARIANT VAL-304.  
 RX MEDLINE=97037613; PubMed=8883260;  
 RA Tanary H., Fromovich Y., Shalom L., Reich Z., Dym O., Lanir N.,  
 RA Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,  
 RA Zaizov R., Seligsohn U.;  
 RA "Ala244Val is a common, probably ancient mutation causing factor VII  
 RT deficiency in Moroccan and Iranian Jews.";  
 RL Thromb. Haemost. 76:283-291(1996).  
 RN [21]  
 RP VARIANTS MALTA THR-194 AND VAL-304.  
 RX MEDLINE=98112461; PubMed=9452082;  
 RA Alshinawi C., Scerri C., Gaidies R., Aquilina A., Felice A.E.;  
 RA "Two new missense mutations (P134T and A244V) in the coagulation  
 RT factor VII gene.";  
 RL Hum. Mutat. Suppl. 1:S189-S191(1998).  
 CC -1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZMOGEN FORM. FACTOR VII IS  
 CC CONVERTED TO FACTOR VIIA BY FACTOR XIa, FACTOR XIIa, FACTOR IXa, OR  
 CC THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR  
 CC AND CALCIUM IONS, FACTOR VIIa THEN CONVERTS FACTOR X TO FACTOR Xa  
 CC BY LIMITED PROTEOLYSIS. FACTOR VIIa WILL ALSO CONVERT FACTOR IX TO  
 CC FACTOR IXa IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to  
 CC form factor Xa.  
 CC -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED  
 CC BY A DISULFIDE BOND.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: PLASMA.  
 CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -1- DISEASE: DEFECTS IN F7 CAN CAUSE COAGULOPATHY.  
 CC -1- PHARMACEUTICAL: Available under the names Niasase or Novoseven  
 CC (Novo Nordisk). Used for the treatment of bleeding episodes in  
 Alignment Scores: 1.33e-34 Length: 466  
 Pred. No.:

Score:	783.00	Matches:	164
Percent Similarity:	56.74%	Conservative:	76
Best Local Similarity:	38.77%	Mismatches:	147
Query Match:	33.50%	Indels:	36
DB:	1	Gaps:	10
US-09-763-153-1 (1-1245) x FA7_HUMAN (1-466)			
QY	1	GCCAACTCCCTCTGGAGGAGTCCCTGACAGCGCTGGAGCGGAGTGCATAGGAG	60
DB	61	AlaAsnAlaPheLeuGluLeuArgProGlySerLeuGluArgGlnCysLeuGlu	80
QY	61	ATCTGTCACATCGAGGAGGCAAGAAATTTCCAAATTTGTCATGACACTGCCTTC	120
QY	81	GlnCysSerPheGluGluAlaArgGluPheLysAspAlaGluArgThrLysLeuPhe	100
DB	121	TGGTCCAAAGCAGCTGCGAGTGCACAGTGTGCTGCTGCTGCTGCGACCCGCTGC	180
DB	101	TrpLysSerTyrSerAspGlyAspGlnCys	112
QY	181	AGCTGTGCTCGGCGCAGCGCTGCTGACAGCGCATCGGCGCTGCTGCTGCTGCTG	240
DB	113	SerProCysGlnAsnGlyGlySerCysLysAspGlnLeuGlnSerTyrIleCysPheCys	132
QY	241	CGCAGCGCTGGAGGCGCTTCTGCCAG---CGCAGGTGAGCTTCCTCAATTGCTCG	297
DB	133	LeuProAlaPheGluGluArgAsnCysGluThrHisLysAspAspGlnLeuIleCysVal	152
QY	298	CTGACACAGCGGCTGACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	354
DB	153	AsnGluAsnGlyGlyCysGluGlnTyrSerSerHisThrGlyThrLysArgSerCys	172
QY	355	AGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	414
DB	173	ArgCysHisGluGlyTyrSerLeuLeuAlaAspGlyValSerCysThrProThrValGlu	192
QY	415	TTCCCTGTGTGGAGCGCTTGGAGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAG	474
DB	193	TyrProCysGlyLys---IleProIleLeuGluLysArgAsnAla	206
QY	475	ACAGAGACCAAGACCAAGTAGATCGCGCTCATTGATGGAAGATGACCGCGCG	534
DB	207	-----SerLysProGlnGlyArgIleValGlyLysValCysProLys	221
QY	535	GGAGACAGCCCTGGAGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	594
DB	222	GlyGlyCysProIleValLeuLeuLeuValAsnGlyAlaGlnLeu---CysGlyGly	240
QY	595	GTGCTATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	651
DB	241	ThrLeuIleAsnThrIleTrpValValSerAlaAlaHisCysPheAspLysLysAsn	260
QY	652	-----AAGTCTCTGTGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	705
DB	261	TrpArgAsnLeuIleAlaValLeuGlyHisAspLeuSerGluHisAspGlyAspGlu	280
QY	706	CTGACCTGACATCAAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	765
DB	281	GlnSerArgArgValAlaGlnValIleIleProSerThrTyrValProGlyThrAsn	300
QY	766	ATGATCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	825
DB	301	HisAspIleAlaLeuLeuArgLeuHisGlnProValValLeuThrAspHisValPro	320
QY	826	ATCTGCTCCGCGAGCGGCTTGTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG	885
DB	321	LeuCysLeuProGluArgThrPheSerGluArgThrLeuAlaPheVal---ArgPheSer	339
QY	886	CTGCTGACGGCTGGGCTACACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG	945
DB	340	LeuValSerGlyTrpGlnLeuLeuAspArgGlyAlaThrAla	354
QY	946	TTGCTCTCACTTCATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1005

DB	355	LeuGluLeuMetValLeuAsnValProLeuMetThrGlnAspCysLeuGlnGlnSer	374
QY	1006	-----AGCAACATGGTCTGAGACATGCTGTGGCGGCGCATCTCTCGG	1050
DB	375	ArgLysValGlyAspSerProAsnIleThrGluTyrMetPheCysAlaGlyTyrSerAsp	394
QY	1051	GACCGGAGGATCGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1110
DB	395	GlySerLysAspSerCysGlyAspSerGlyGlyProHisAlaThrHisTyrArgGly	414
QY	1111	ACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1170
DB	415	ThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThrValGlyHisPhe	434
QY	1171	GCGTTCACCAACAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1230
DB	435	GlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeuMetArgSerGlu	454
QY	1231	GAAGCGCGCC	1239
DB	455	ProArgPro	457
RESULT 14			
FA7_BOVIN STANDARD; PRT; 407 AA.			
ID	FA7_BOVIN	STANDARD;	PRT; 407 AA.
AC	P22457;		
DT	01-AUG-1991 (Rel. 19, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Coagulation factor VII (EC 3.4.21.21) (Serum prothrombin conversion accelerator).		
GN	P7.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=8900362; PubMed=3049594;		
RA	Takeya H., Kawabata S., Nakagawa K., Yamamichi Y., Miyata T.,		
RA	Iwanaga S.;		
RT	"Bovine factor VII. Its purification and complete amino acid		
RT	sequence."		
RL	J. Biol. Chem. 263:14868-14877(1988).		
RN	[2]		
RP	STRUCTURE OF CARBOHYDRATE ON SER-52.		
RX	MEDLINE=8921399; PubMed=3149637;		
RA	Hase S., Kawabata S., Nishimura H., Takeya H., Sueyoshi T.,		
RA	Miyata T., Iwanaga S., Takao T., Shimomishi Y., Ikenaka T.;		
RT	"A new trisaccharide sugar chain linked to a serine residue in bovine		
RT	blood coagulation factors VII and IX."		
RL	J. Biochem. 104:867-868(1988).		
RN	[3]		
RP	STRUCTURE OF CARBOHYDRATE ON SER-52.		
RX	MEDLINE=91344709; PubMed=2129367;		
RA	Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;		
RT	"A new trisaccharide sugar chain linked to a serine residue in the		
RT	first EGF-like domain of clotting factors VII and IX and protein Z."		
RL	Adv. Exp. Med. Biol. 281:121-131(1990).		
CC	-1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS		
CC	CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XI, FACTOR XII, OR		
CC	THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR		
CC	AND CALCIUM IONS, FACTOR VIIA THEN CONVERTS FACTOR X TO FACTOR IXA		
CC	BY LIMITED PROTEOLYSIS. FACTOR VIIA WILL ALSO CONVERT FACTOR IX TO		
CC	FACTOR IXA IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.		
CC	-1- CATALYTIC ACTIVITY: Hydrolyzes one Arg- -Ile bond in factor X to		
CC	form factor Xa.		
CC	-1- SUBUNIT: HETERO-DIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED		
CC	BY A DISULFIDE BOND.		
CC	-1- TISSUE SPECIFICITY: PLASMA.		
CC	-1- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME		

CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC PIR: A31979; A31979.  
 DR HSP: P08709; Ibf9.  
 DR MEROPS: S01.215; .  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00010; EGF\_blood.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_CA; 1.  
 DR SMART: SM00001; EGF\_like; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase: Serine protease; Blood coagulation; Zymogen; Glycoprotein;  
 KW Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;  
 KW EGF-like domain; Repeat.  
 FT CHAIN 1 152  
 FT DOMAIN 153 407  
 FT CHAIN 153 407  
 FT DOMAIN 6 35  
 FT CHAIN 46 82  
 FT DOMAIN 87 128  
 FT CHAIN 153 407  
 FT DOMAIN 152 153  
 FT SITE 152 153  
 FT ACT\_SITE 193  
 FT ACT\_SITE 242  
 FT BINDING 344  
 FT BINDING 338  
 FT DISULFID 17 22  
 FT DISULFID 50 61  
 FT DISULFID 55 70  
 FT DISULFID 72 81  
 FT DISULFID 91 102  
 FT DISULFID 98 112  
 FT DISULFID 114 127  
 FT DISULFID 135 162  
 FT DISULFID 159 164  
 FT DISULFID 178 194  
 FT DISULFID 310 329  
 FT DISULFID 340 368  
 FT MOD\_RES 6 7  
 FT MOD\_RES 7 6  
 FT MOD\_RES 14 14  
 FT MOD\_RES 16 16  
 FT MOD\_RES 19 19  
 FT MOD\_RES 20 20  
 FT MOD\_RES 25 25  
 FT MOD\_RES 26 26  
 FT MOD\_RES 29 29  
 FT MOD\_RES 35 35  
 FT MOD\_RES 52 52  
 FT CARBOHYD 52 52  
 FT CARBOHYD 145 145

GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CALCIUM.  
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 PIR: A31979; A31979.  
 HSP: P08709; Ibf9.  
 MEROPS: S01.215; .  
 InterPro: IPR000152; Asx\_hydroxyl.  
 InterPro: IPR001314; Chymotrypsin.  
 InterPro: IPR000561; EGF-like.  
 InterPro: IPR000742; EGF-2.  
 InterPro: IPR001881; EGF-Ca.  
 InterPro: IPR001438; EGF-II.  
 InterPro: IPR002383; GLA\_blood.  
 InterPro: IPR001254; Ser\_protease\_Try.  
 Pfam: PF00008; EGF\_2.  
 Pfam: PF00089; trypsin; 1.  
 Pfam: PF00594; gla; 1.  
 PRINTS: PR00722; CHYMOTRYPSIN.  
 PRINTS: PR00010; EGF\_blood.  
 PRINTS: PR00001; GLABLOOD.  
 SMART: SM00179; EGF\_CA; 1.  
 SMART: SM00001; EGF\_like; 1.  
 SMART: SM00069; GLA; 1.  
 SMART: SM00020; Tryp\_Spc; 1.  
 PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 PROSITE: PS00022; EGF\_1; 1.  
 PROSITE: PS01186; EGF\_2; 2.  
 PROSITE: PS01187; EGF\_CA; 1.  
 PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 PROSITE: PS00135; TRYPSIN\_SER; 1.  
 Hydrolase: Serine protease; Blood coagulation; Zymogen; Glycoprotein;  
 Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;  
 EGF-like domain; Repeat.  
 CHAIN 1 152  
 DOMAIN 153 407  
 CHAIN 153 407  
 DOMAIN 6 35  
 CHAIN 46 82  
 DOMAIN 87 128  
 CHAIN 153 407  
 DOMAIN 152 153  
 SITE 152 153  
 ACT\_SITE 193  
 ACT\_SITE 242  
 BINDING 344  
 BINDING 338  
 DISULFID 17 22  
 DISULFID 50 61  
 DISULFID 55 70  
 DISULFID 72 81  
 DISULFID 91 102  
 DISULFID 98 112  
 DISULFID 114 127  
 DISULFID 135 162  
 DISULFID 159 164  
 DISULFID 178 194  
 DISULFID 310 329  
 DISULFID 340 368  
 MOD\_RES 6 7  
 MOD\_RES 7 6  
 MOD\_RES 14 14  
 MOD\_RES 16 16  
 MOD\_RES 19 19  
 MOD\_RES 20 20  
 MOD\_RES 25 25  
 MOD\_RES 26 26  
 MOD\_RES 29 29  
 MOD\_RES 35 35  
 MOD\_RES 52 52  
 CARBOHYD 52 52  
 CARBOHYD 145 145

FT CARBOHYD 203 203 N-LINKED (GLCNAC...)  
 SQ SEQUENCE 407 AA; 44431 MW; 703E1FE0636F7F10 CRC64;  
 Alignment Scores:  
 Pred. No.: 2,04e-34 Length: 407  
 Score: 779.50 Matches: 166  
 Percent Similarity: 54.89% Conservative: 64  
 Best Local Similarity: 39.62% Mismatches: 150  
 Query Match: 33.35% Indels: 39  
 DB: 1 Gaps: 11  
 US-09-763-153-1 (1-1245) x FAF7\_BOVIN (1-407)  
 QY 1 GCCAATCTCTCTCTGGAGGAGCTCGTCACAGCAGCTGGAGCGGAGTGTGATAGAGGAG 60  
 Db 1 AlaAsnGlyPheLeuGluGluLeuLeuProGlySerLeuGluArgGluCysArgGluGlu 20  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTTCCAAATGTGGATGACACACTGGCCCTTC 120  
 Db 21 LeuCysSerPheGluGluAlaHisGluLeuPheArgAsnGluGluArgThrArgGlnPhe 40  
 QY 121 TGGTCCAAAGCAGCTCGACGGTGCACAGTGTCTTGGCCCTTGGAGCACCCTGGCCG 180  
 Db 41 TrpValSerTyrAsnAspGlyAspGlnCys-----AlaSer 52  
 QY 181 AGCTGTGCTCGGGCAGCGCACGTGATCGAGCGCATCGAGCTTCAGCTTCAGCTCGGACTGC 240  
 Db 53 SerProCysGlnAsnGlyGlySerCysGluAspGlnLeuArgSerTyrIleCysPheCys 72  
 QY 241 CGCAGCGCTGGAGGCGCTCTTCCACAGCGGAG---GTGAGCTTCCTCAATTCCTCG 297  
 Db 73 ProAspGlyPheGluGlyArgAsnCysGluThrAspLysGlnSerGlnLeuIleCysAla 92  
 QY 298 CTGGCAACCGCGCTGCACGCATTTACTGCTAGAGAGGTGGG---TGGCGGCGCTGT 354  
 Db 93 AsnAspAsnGlyGlyCysGluGlnTyrCysGlyAlaAspProGlyAlaGlyArgPheCys 112  
 QY 355 AGCTGTGCTCGCTACAGCTGGGGAGGAGCTCTCGACGTGTACCCCGCTGAGTGAAG 414  
 Db 113 TrpCysHisGluGlyTyrAlaLeuGlnAlaAspGlyValSerCysAlaProThrValGlu 132  
 QY 415 TTCCCTTGGGAGCGCTTGGAGCGGATGGAGAGCGAGCTCACCTGNAACGAGAC 474  
 Db 133 TyrProCysGlyLys---IleProValLeuGluLysArgAsnGly----- 146  
 QY 475 ACAGAGACCAAGACCAAGTAGTACCGCGCTCATGTATGGAGAGATGACACCGCGG 534  
 Db 147 -----SerLysProGlnGlyArgIleValGlyGlyHisValCysProLys 161  
 QY 535 GGAGACAGCCCTTGGCAGGTGCTCTGCTGGACTCAAAGAAGAGCTGGCTGGCGGCA 594  
 Db 162 GlyGluCysProTrpGln---AlaMetLeuLysLeuAsnGlyAlaLeuLeuCysGlyGly 180  
 QY 595 GTGCTCATCCCTCTCTGGGTGCTGACAGCGGCCCTCATGTGATGATGATCAAG--- 651  
 Db 181 ThrLeuValGlyProAlaTrpValValSerAlaAlaHisCysPheGluArgLeuArgSer 200  
 QY 652 -----AAGCTCTTGTGAGCTTGGAGATGATGACCTGCGCGCTGGGAGAGTGGGAG 705  
 Db 201 ArgGlyAsnLeuThrAlaValLeuGlyLysAspLeuSerArgValGluGlyProGlu 220  
 QY 706 CTGGACCTGGACATCAGGAGGTCTTGTCCACCCCACTACAGAGAGCACCACCGAC 765  
 Db 221 GlnGluArgArgValAlaGlnIleValProLysGlnTyrValProGlyGlnThrAsp 240  
 QY 766 AATGACATCGACCTGCTGACCTGGCCCGCCGCGCCCTCTCGAGACCATAGTGGCC 825  
 Db 241 HisAspValAlaLeuLeuGlnLeuAlaGlnProValAlaLeuGlyAspHisValAlaPro 260  
 QY 826 ATCTGCTCCCGGACAGCGGCTTGCAGAGCGAGCTCAATCAGCGCGCCGAGGAGACC 885  
 Db 261 LeuCysLeuProAspProAspPheAlaAspGlnThrLeuAlaPheVal---ArgPheSer 279



Score:	770.00	Matches:
Percent Similarity:	55.71%	Conservative:
Best Local Similarity:	38.81%	Mismatches:
Query Match:	32.95%	Indels:
DB:	1	Gaps:
US-09-763-153-1 (1-1245) x FA7_MOUSE (1-446)		
QY	1	GCCAACTCTTCCTGGAGGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60
Db	42	AlaAsnSerLeuLeuGluLeuLeuTprGlySerLeuGluArgGlyCysAsnGluGlu 61
QY	61	ATCTGTGACTTCAGAGAGCCACAGGAATTTCCAAATATGTGGATGACACACTGGCCCTTC 121
Db	62	GlnCysSerPheGluAlaArgGluLeuPheLeuPheLeuSerProGluArgThrLysGlnPhe 81
QY	121	TGTTCCNAGCACTGCAGCGGTGACCACTGCTGGTCTTGCCCTTGGAGCACCCCGTGGCC 181
Db	82	TrpIleValTyrSerAspGlyAspGlnCysAlaSerAsnPro----- 95
QY	181	AGCGTGTGCTGGCGGCACGCGCATGTCAGCGGCATCGGAGCTTCAGTTCGAGTGC 240
Db	96	-----CysGlnAsnGlyClyThrCysGlnAspHisLeuLysSerTyrValCysPheCys 113
QY	241	CGCAGCGCTGGAGGGCGCTTCTCCAGCGCGAGGTGAGC---TTCTCAATGTCTCG 297
Db	114	LeuLeuAspPheGluGlyArgAsnCysGlnLysSerLysAsnGluGlnLeuLeuLeuCysAla 133
QY	298	CTGGACAACGCGGCTGCAGCGCATPACTGCTTACGAGAGGTGGCTGGCGGCGC---TGT 354
Db	134	AsnGluAsnGlyAspCysAspGlnTyrCysArgAspHisValGlyThrLysArgThrCys 153
QY	355	AGCTGTGCGCTGGCTACAAGCTGGGGGACACCTCTGCAGTCTCACCCCGCAGTGAAG 414
Db	154	SerCysHisGluAspTyrThrLeuGlnProAspGluValSerCysLysProLysValGlu 173
QY	415	TTCCCTTGTGGGAGCCCTGGAAGCGGATGGAGAAGCGCAGTCACTGCTGAACAGCAGAC 474
Db	174	TyrProCysGlyArg---IleProValValGluLysArgAsnSerSerArgGlnGly 192
QY	475	ACAGAAGACCAGAAGACCAAGTAGTACCGCGCTATTGATCGGAAGATGACCGGCGG 534
Db	193	-----ArgIleValGlyAsnValCysProLys 202
QY	535	GGAGACACCCCTGGCAGGTGGTCTGCTGGACTCAAGAAGAAGCTGCCTCGCGGCGCA 594
Db	203	GlyGluCysProTrpGlnAlaVal---LeuLysIleAsnGlyLeuLeuLeuLysGlyAla 221
QY	595	GTGCTCATCCACCCTCTCGGTGTCTGACAGCGGCCACTGCTGATGGATGAGTCCCAAG--- 651
Db	222	ValLeuLeuAspAlaAspGlyTrpIleValThrAlaAlaHisCysPheAspAsnIleArgTyr 241
QY	652	-----AAGTCTCTTGCAGGCTTGGAGAGTATGACCTCGCGCGCTGGGGAAGTGGGAG 705
Db	242	TrpGlyAsnIleThrValValMetGlyGluHisAspPheSerGluLysAspGlyAspGlu 261
QY	706	CTGGACCTGGACATCAAGGAGGTCTTCGTCACCCCACTACAGCAAGAGCACCACCGAC 765
Db	262	GlnValArgValThrGlnValIleMetProAspLysTyrIleArgGlyLysIleAsn 281
QY	766	AATGACATCCACTGTGCACCTGGCCCCAGCCGCCACCTCTCCAGACCATAGTGGCC 825
Db	282	HisAspIleAlaLeuLeuArgLeuHisArgProValThrPheThrAspTyrValValPro 301
QY	826	ATCTGCTCCCGACAGCGGCCTTCAGAGCGCGAGCTCAATCATCGCGCGCCAGGAGACC 885
Db	302	LeuCysLeuProGlnLysSerPheSerGluAsnThrLeuAlaArgIle---ArgPheSer 320
QY	886	CTCGTGACGGCTGGGGCTTACCACAGCAGCCGACAGAGAAGGAGGCGCAAGAGAACCGCACC 945
Db	321	ArgValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla----- 335
QY	946	TTCTGCTCTCAACTTCATCAAGATTCCCTGGTCCCGCACAATAGTGCACGAGGTGTCATG 1005







Cell 45:343-348(1986).

[14] STRUCTURE OF CARBOHYDRATE ON SER-99.  
MEDLINE=90062160; PubMed=2511202;  
Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T., Takao T., Shimonishi Y., Iwanaga S.;  
RA "Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl2-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z.";  
RT J. Biol. Chem. 264:20320-20325(1989).  
[15] STRUCTURE OF CARBOHYDRATE ON SER-99.  
MEDLINE=91344709; PubMed=2129367;  
RA Iwanaga S., Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T.;  
RA "A new trisaccharide sugar chain linked to a serine residue in the first EGF-like domain of clotting factors VII and IX and protein Z.";  
RT Adv. Exp. Med. Biol. 281:121-131(1990).  
[16] STRUCTURE OF CARBOHYDRATE ON SER-107.  
MEDLINE=92388094; PubMed=1517205;  
RA Nishimura H., Takao T., Hase S., Shimonishi Y., Iwanaga S.;  
RA "Human factor IX has a tetrasaccharide O-glycosidically linked to serine 61 through the fucose residue.";  
RT J. Biol. Chem. 267:17520-17525(1992).  
[17] PHOSPHORYLATION OF SER-114.  
Harris R.J., Papac D.I., Truong L., Smith K.J.;  
RA "Partial phosphorylation of serine-68 in EGF-1 of human factor IX.";  
RT (In) Abstracts of Xth International conference on methods in protein structure analysis, pp.50-50, Annecy (1996).  
[18] POST-TRANSLATIONAL MODIFICATIONS.  
MEDLINE=20575397; PubMed=11133752;  
RA Arruda V.R., Hagstrom J.N., Deitch J., Heiman-Patterson T., Camire R.M., Chu K., Fields P.A., Herzog R.W., Couto L.B., Larson P.J., High K.A.;  
RA "Posttranslational modifications of recombinant myotube-synthesized human factor IX.";  
RT Blood 97:130-138(2001).  
[19] STRUCTURE BY NMR OF 47-93.  
MEDLINE=95229607; PubMed=7713897;  
RA Freedman S.J., Furie B.C., Furie B., Baleja J.D.;  
RA "Structure of the metal-free gamma-carboxyglutamic acid-rich membrane binding region of factor IX by two-dimensional NMR spectroscopy.";  
RT J. Biol. Chem. 270:7980-7987(1995).  
[20] STRUCTURE BY NMR OF 47-93.  
MEDLINE=96032604; PubMed=7547952;  
RA Freedman S.J., Furie B.C., Furie B., Baleja J.D.;  
RA "Structure of the calcium ion-bound gamma-carboxyglutamic acid-rich domain of factor IX.";  
RT Biochemistry 34:12126-12137(1995).  
[21] STRUCTURE BY NMR OF 47-93.  
MEDLINE=96279169; PubMed=8663165;  
RA Freedman S.J., Bluestein M.D., Baleja J.D., Jacobs M., Furie B.C., Furie B.;  
RA "Identification of the phospholipid binding site in the vitamin K-dependent blood coagulation protein factor IX.";  
RT J. Biol. Chem. 271:16227-16236(1996).  
[22] STRUCTURE BY NMR OF 47-93.  
MEDLINE=97199336; PubMed=9047312;  
RA Li L., Darden T.A., Freedman S.J., Furie B.C., Furie B., Baleja J.D., Smith H., Hiskey R.G., Pedersen L.G.;  
RA "Refinement of the NMR solution structure of the gamma-carboxyglutamic acid domain of coagulation factor IX using molecular dynamics simulation with initial Ca2+ positions determined by a genetic algorithm.";  
RT Biochemistry 36:2132-2136(1997).  
[23] STRUCTURE BY NMR OF 91-133.  
MEDLINE=91308127; PubMed=1854745;  
RA Huang L.H., Cheng H., Pardi A., Tam J.P., Sweeney W.V.;  
RA "Sequence-specific 1H NMR assignments, secondary structure, and location of the calcium binding site in the first epidermal growth factor like domain of blood coagulation factor IX.";  
RT Biochemistry 30:7402-7409(1991).  
[24] STRUCTURE BY NMR OF 92-130.  
MEDLINE=92284090; PubMed=1304885;  
RA Baron M., Norman D.G., Harvey T.S., Handford P.A., Mayhew M., Tse A.G.D., Brownlee G.G., Campbell I.D.C.;  
RA "The three-dimensional structure of the first EGF-like module of human factor IX: comparison with EGF and TGF-alpha.";  
RT Protein Sci. 1:81-90(1992).  
[25] X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 92-130.  
MEDLINE=95330802; PubMed=7606779;  
RA Rao Z., Handford P., Mayhew M., Knott V., Brownlee G.G., Stuart D.;  
RA "The structure of a Ca(2+)-binding epidermal growth factor-like domain: its role in protein-protein interactions.";  
RT Cell 82:131-141(1995).  
[26] MOLECULAR PATHOLOGY OF HEMOPHILIA B.  
US-09-763-153-1 (1-1245) x FA9\_HUMAN (1-461)  
QY 13 CTGAGGAGGCTCCGTACACAGCGGTGGAGTGCATAGAGGATCTGTGACTTC 72  
Db 52 LeuGluGluPheValGlnGlyAsnLeuGluGluGlyCysMetGluGluGlyCysSerPhe 71  
QY 73 GAGGAGGCCAAGGAAATTTCCAAATGTGGATGCACACTGGCCCTTGTGTCACAGC 132  
Db 72 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheThrPheGlnTyr 91  
QY 133 GTCGACGGTGACAGTGTGCTTGGCTTGGCCCTTGGAGCACCCTGCGCTGTGTGTC 192  
Db 92 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 103  
QY 193 GGGCAGCGCAGCTGCATCGACGGCATCGGACGCTTACGCTGCAGTCCGCGAGCGGTGG 252  
Db 104 AsnGlyGlySerCysLysAspAspLeuAsnSerTyrGluCysTyrProPheGlyPhe 123  
QY 253 GAGGCGCGCTTGTGCACGGCGAGGTGAGTCTCTCAATTCCTGCTGGACACGCGCGC 312  
Db 124 GluGlyLysAsnGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 140  
QY 313 TGCACGCATTACTGCTAGAGGAGGTGGCTGGCGG---CGCTGTAGCTGTGCGCTGGC 369  
Db 141 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 160  
QY 370 TACAAGCTGGGGACGACCTCTCTGAGTGTACCCCGCAGTCAAGTTCCTGTGGGAGG 429  
Db 161 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 180  
QY 430 CCCTGGAAGCGGATGGAGAAGAGCGACCTACCTGAAACGA----- 471  
Db 181 ValSerValSerGlnThrSerLysLeuThrArgAlaGluAlaValPheProAspValAsp 200  
QY 472 -----GACACAGAGACCAAGACCAAGTAGAT----- 501  
Db 201 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer 220  
QY 502 -----CCGGGGCTCATTTGATGGAGATCACCAGCGGGGACAGCCCTGG 549

Db 221 PheAsnAspPheThrArgValValGlyGlyAlaLysProGlyGlnPheProTrp 240  
 Qy 550 CAGGTGGTCTGCTGGACTCAAGAGAGAGCTGGCTGGGGGAGTGTCTATCCACCCC 609  
 Db 241 GlnValVal---LeuAsnGlyLysValAlaPheCysGlySerIleValAsnGlu 259  
 Qy 610 TCCTGGGTGTGACAGCGGCCACTGCATGGATGAGTCAACAAGCTCTGTGTGAGGCTT 669  
 Db 260 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla 279  
 Qy 670 GGAGATGATGACCTGGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGGAGGTC 729  
 Db 280 GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 299  
 Qy 730 TTCGTCCACCCCACTACAGCAGCAGCACCACCC-----GACATGACATCGCTGCT 783  
 Db 300 IleProHisHisAsnTyrAsnAlaIleAsnLysTyrAsnHisAspIleAlaLeu 319  
 Qy 784 CACCTGGCCCGCCCGCCACCTCTCCGACACCATAGTGCCTGCTCCCGACAGC 843  
 Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 339  
 Qy 844 GGCCTTGACAGCGGAGCTCAATCAGCGCGCCAGGAGACCTCGTACGGGCTGGGC 903  
 Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 357  
 Qy 904 -----TACCACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957  
 Db 358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370  
 Qy 958 TTCATCAAGATCCCGTGGTCCCGCACATCAGTGCAGCAGGATCATGAGCAACATGGTG 1017  
 Db 371 TyrLeuArgValProLeuValAlaPheAlaThrCysLeuArgSerThrLysPheThrIle 390  
 Qy 1018 TCTGAGACATGCTGTGTGGGGCATCTCCGGGACCGGAGGATCGCTCGGAGGCGAC 1077  
 Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410  
 Qy 1078 AGTGGGGGGCCATGCTGCTCTCTCCAGCGCACCTGGTCTGTGGTGGCTGTGTGAGC 1137  
 Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430  
 Qy 1138 TGGGGTGGGCTGTGGCTCTTCAACACTAGCGGCTTTACACCAAGTACAGCGCTAC 1197  
 Db 431 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleThrLysValSerArgTyr 450  
 Qy 1198 CTCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454  
 RESULT 18  
 FA9\_MOUSE STANDARD; PRT; 459 AA.  
 P16294;  
 01-AUG-1990 (Rel. 15, Created)  
 01-FEB-1991 (Rel. 17, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor)  
 DE (Fragment).  
 GN F9 OR CF9.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90215309; PubMed=2323576;  
 RA Wu S.-M., Stafford D.W., Ware J.;  
 RT "Deduced amino acid sequence of mouse blood-coagulation factor IX.";  
 RL gene 86:275-278(1990).  
 RN [2]  
 RP SEQUENCE OF 168-451 FROM N.A.  
 RX MEDLINE=90152675; PubMed=2303254;

RA Sarker G., Koerberl D.D., Sommer S.S.;  
 RT "Direct sequencing of the activation peptide and the catalytic domain  
 of the factor IX gene in six species.";  
 RL Genomics 6:133-143(1990).  
 CC -1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT  
 CC PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY  
 CC CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++  
 CC IONS, PHOSPHOLIPIDS, AND FACTOR VIIIa.  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to  
 CC form factor Xa.  
 CC -1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE  
 CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2  
 CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M23109; AAA37629.1; -;  
 CC EMBL; M26236; AAA37630.1; -;  
 CC PIR; J00419; J00419.  
 CC HSSP; P00740; ICFH.  
 CC MEROPS; S01.214; -;  
 CC MGD; MGI-88384; F9.  
 CC InterPro; IPR000152; Asx\_hydroxyl.  
 CC InterPro; IPR000561; EGF-like.  
 CC InterPro; IPR000742; EGF-2.  
 CC InterPro; IPR001881; EGF-Ca.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC InterPro; IPR000294; VitK\_dep\_GLA.  
 CC Pfam; PF00008; EGF\_2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00594; g1a; 1.  
 CC SMART; SM00179; EGF\_CA; 1.  
 CC SMART; SM00069; GLA; 1.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; 2.  
 CC PROSITE; PS01187; EGF\_CA; 1.  
 CC PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC KW Blood coagulation; Plasma; Serine protease; Calcium-binding;  
 CC KW Hemolysis; Glycoprotein; Vitamin K; Gamma-carboxyglutamic acid;  
 CC KW Hemophilia; Hydroxylation; Zymogen; Signal; EGF-like domain;  
 CC Repeat.  
 CC FT NON-TER 1 1  
 CC FT SIGNAL <1 16  
 CC FT PROPEP 17 34  
 CC FT CHAIN 35 180  
 CC FT PROPEP 181 224  
 CC FT CHAIN 225 459  
 CC FT DOMAIN 81 117  
 CC FT DOMAIN 118 159  
 CC FT DOMAIN 225 459  
 CC FT SITE 180 181  
 CC FT SITE 224 225  
 CC FT MOD\_RES 41 41  
 CC FT MOD\_RES 42 42  
 CC FT MOD\_RES 49 49  
 CC FT MOD\_RES 51 51  
 CC FT  
 CC BY SIMILARITY.  
 CC FACTOR IXA (ACTIVE FORM) LIGHT CHAIN.  
 CC ACTIVATION PEPTIDE.  
 CC FACTOR IXA (ACTIVE FORM) HEAVY CHAIN.  
 CC EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 CC EGF-LIKE 2.  
 CC SERINE PROTEASE.  
 CC CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).  
 CC CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).  
 CC GAMMA-CARBOXYGLUTAMIC ACID  
 CC (BY SIMILARITY).  
 CC GAMMA-CARBOXYGLUTAMIC ACID  
 CC (BY SIMILARITY).  
 CC GAMMA-CARBOXYGLUTAMIC ACID  
 CC (BY SIMILARITY).  
 CC GAMMA-CARBOXYGLUTAMIC ACID  
 CC (BY SIMILARITY).  
 CC GAMMA-CARBOXYGLUTAMIC ACID  
 CC (BY SIMILARITY).

```

FT MOD_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 61 61 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 67 67 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 74 74 GAMMA-CARBOXYGLUTAMIC ACID
FT FT (BY SIMILARITY)
FT MOD_RES 98 98 HYDROXYLATION (BY SIMILARITY)
FT ACT_SITE 265 265 CHARGE RELAY SYSTEM.
FT ACT_SITE 313 313 CHARGE RELAY SYSTEM.
FT ACT_SITE 409 409 CHARGE RELAY SYSTEM.
FT DISULFID 52 57 BY SIMILARITY.
FT DISULFID 85 96 BY SIMILARITY.
FT DISULFID 90 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 122 133 BY SIMILARITY.
FT DISULFID 129 143 BY SIMILARITY.
FT DISULFID 145 158 BY SIMILARITY.
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 363 363 H -> Q (IN REF. 2).
FT CONFLICT 388 388 T -> I (IN REF. 2).
SQ SEQUENCE 459 AA; 51635 MW; EF439C84D06BC1A CRC64;

Alignment Scores:
Pred. No.: 1,36e-31 Length: 459
Score: 726.00 Matches: 155
Percent Similarity: 50.91% Conservative: 68
Best Local Similarity: 35.39% Mismatches: 151
Query Match: 31.07% Indels: 64
DB: 1 Gaps: 12

US-09-763-153-1 (1-1245) x FA9_MOUSE (1-459)

QY 13 CTGAGGAGCTCCGTCACAGCAGCCCTGGAGCGGAGTGATGATGAGAGAGATCTGTGACTTC 72
|||||
40 LeuGluGluPheValArgGlyAsnLeuGluArgGluCysIleGluGluArgCysSerPhe 59
|||||
73 GAGGAGCCAGGAATTTCCAAATGTTGGATGACACACTGGCCTTCTGGTCCCAAGCAC 132
|||||
60 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheTrpLysGlnTyr 79
|||||
133 GTCGACGCTGACAGTCTGGTCTTCCCTTGGAGCACCCTGCGCCAGCCTGTGCTGC 192
|||||
80 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 91
193 GGGCAGCGCAGTCATCGACGCGATCGGAGCTTCAGTGCAGCTGCGCGAGCGCTGG 252
|||||
92 AsnGlyGlyIleCysLysAspAspIleSerSerTyrGluCysTrpCysGlnValGlyPhe 111
253 GAGGCGCGCTTCTCCAGCGGAGGTGAGTCTCTCAATGCTGCTGGGACACGCGCGC 312
|||||
112 GluGlyArgAsnCysGluLeuAspAlaThr-----CysAsnIleLysAsnGlyArg 128
313 TGCAGCGCATTAAGTCTAGAGAGGTGGCTGGCGG---CGCTGAGCTGTGCGCTGGC 369
129 CysLysGlnPheCysLysAsnSerProAspAsnLysValIleCysSerCysThrGluGly 148
370 TACAGCTGGGAGCAGCTCCTCGAGTGTCACCCCGCAGTGAAGTCCCTTGTGGAGG 429
149 TyrGlnLeuAlaGluAspGlnLysSerCysGluProThrValProPheProCysGlyArg 168
430 CCC-----TGCAAGCGGATG 444

```

```

Db 169 AlaSerIleSerTyrSerSerLysLysIleThrArgAlaGluThrValPheSerAsnMet 188
QY 445 GAGAAGAGCGCAGTCAC-----
Db 189 AspTyrGluAsnSerThrGluAlaValPheIleGlnAspAspIleThrAspGlyAlaIle 208
QY 463 CTGAACGAGACACAGACAGACCAAGAACCAAGTAGAT--CGCGGCTCAATTGATGG 519
Db 209 LeuAsnAsnValThrGluSerSerGluSerLeuAsnAspPheThrArgValValGlyGly 228
QY 520 AAGATGACAGCGGCGGAGACAGCCCTGCGAGGTGCTCTGCTGAGTCAAGAGAGAAG 579
Db 229 GluAsnAlaLysProGlyIleProTrpGlnValIle---LeuAsnGlyGluIleGlu 247
QY 580 CTGGCTGCGGGGAGTCTCATCCACCCCTCTGGTGGTGTGACAGCGGCGCCACTGCATG 639
Db 248 AlaPheCysGlyGlyAlaIleAsnGluLysTrpIleValThrAlaAlaHisCysLeu 267
QY 640 GATGACTCCAAAGCTCTTGTGAGGCTTGAGAGATGATGACCTGCGCGCTGGGAGAAG 699
Db 268 LysProGlyAspLysIleGluValAlaGlyGluTyrAsnIleAspLysLysGluAsp 287
QY 700 TGGGAGCTGGACCTGGACATCAAGGAGTCTTCTGTCACCCCACTACAGCAAGAGACC 759
Db 288 ThrGluGlnArgAsnValIleArgThrIleProHisHisGlnTyrAsnAlaThrIle 307
QY 760 ACC-----GACATGACATCCACTGCTGACCTGGCCAGCCGCGCCCTCTCCGAG 813
Db 308 AsnLysTyrSerHisAspIleAlaLeuGluLeuAspLysProLeuIleLeuAsnSer 327
QY 814 ACCATAGTGCATCTGCTCCCGGACAGCGGCTTGGAGAGCGGAGCTCAATCAG--- 870
Db 328 TyrValThrProLysCys-----ValAlaAsnArgGluTyrThrAsnIle 342
QY 871 -----GCCGCGCAGGAGACCCCTGCTGAGCGGCTGGGCTTACCACAGCAGC 915
Db 343 PheLeuLysPheGlySerGly-----TyrValSerGlyTrpGlyLysValPheAsn 359
QY 916 CGAGAGAGGAGGCGCAAGAGAACCGACCTCTGCTCAACTTCATCAAGATTCCCTGTG 975
Db 360 LysGlyArgHisAlaSer-----IleLeuGlnTyrLeuArgValProLeu 374
QY 976 GTCCCGCACAAATGAGTGCAGCGAGGTCTATGACCAACATGGTGTCTGAGAACATGTGTGT 1035
Db 375 ValAspArgAlaThrCysLeuArgSerThrThrPheThrThrTyrAsnAsnMetPheCys 394
QY 1036 GCGGCGCATCTCGGGACCGCGCAGGATGCTGCGAGGCGGACAGTGGGGGCGCCATGTC 1095
Db 395 AlaGlyTyrArgGluGlyGlyLysAspSerCysGluGlyAspSerGlyGlyProHisVal 414
QY 1096 GCCTCTTCCACGCGCAGCTGGTCTGGTGGCTGCTGAGTGGGCTGAGGCGCTGTGGG 1155
Db 415 ThrGluValGluGlyThrSerPheLeuThrGlyIleIleSerTrpGlyGluGluCysAla 434
QY 1156 CTCCTTCCAACTACGCGGTTTACACCAAGTACGCGCTACCTGAGCTGGATC 1209
Db 435 MetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyrValAsnTrpIle 452
RESULT 19
FA10_TROCA STANDARD; PRT; 376 AA.
ID FA10_TROCA
AC P81428;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coagulation factor X (EC 3.4.21.6) (Trocarin prothrombin activator).
OS Tropicdehis carinatus (Australian rough-scaled snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Notechidae; Tropicdehis.
OX NCBI_taxid=100989;
RN [1]

```

SEQUENCE, AND CHARACTERIZATION.

RP TISSUE-Venom;

RX MEDLINE=99326314; PubMed=10397729;

RA Joseph J.S., Chung M.C.M., Jeyaseelan K., Kini R.M.;

RT "Amino acid sequence of trocarn, a prothrombin activator from

RT Tropidochelis carinatus venom: its structural similarity to coagulation

RT factor Xa.;"

RL Blood 94:621-631(1999).

CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that

CC converts prothrombin to thrombin in the presence of factor Va,

CC calcium and phospholipid during blood clotting. Acts as a toxin

CC in venom.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then

CC Arg-|-Ile bonds in prothrombin to form thrombin.

CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR

CC AND ARE HELD TOGETHER BY 1 OR MORE DISULFIDE BONDS.

CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME

CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND

CC CALCIUM.

CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE

CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.

CC HSSP: P00742; IHCG.

CC MEROPS: S01.216; ..

DR InterPro: IPR000152; Asx\_hydroxyl.

DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000742; EGF-2.

DR InterPro: IPR001881; EGF-Ca.

DR InterPro: IPR001438; EGF-II.

DR InterPro: IPR002383; GLA\_blood.

DR InterPro: IPR001254; Ser\_protease\_Try.

DR InterPro: IPR000294; VitK\_dep\_GLA.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR PRINTS: PR00010; EGF\_BLOOD.

DR PRINTS: PR00001; GLABLOOD.

DR SMART: SM00179; EGF\_CA; 1.

DR SMART: SM00001; EGF-like; 1.

DR SMART: SM00069; GLA; 1.

DR SMART: SM00020; Tryp\_Spc; 1.

DR PROSITE: PS00010; ASX\_HYDROXYL; 1.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS01187; EGF\_CA; 1.

DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.

DR PROSITE: PS00240; TRYPSIN\_DOM; 1.

DR PROSITE: PS00134; TRYPSIN\_HIS; 1.

DR PROSITE: PS00135; TRYPSIN\_SER; 1.

KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;

KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;

KW Zymogen; EGF-like domain; Repeat.

CHAIN 1 141

NON\_CONS 141 142

CHAIN 142 376

CHAIN 142 376

DOMAIN 50 81

DOMAIN 89 124

DOMAIN 142 376

MOD\_RES 6 6

MOD\_RES 7 7

MOD\_RES 14 14

MOD\_RES 16 16

MOD\_RES 19 19

MOD\_RES 20 20

MOD\_RES 25 25

MOD\_RES 26 26

MOD\_RES 29 29

MOD\_RES 32 32

MOD\_RES 35 35

MOD\_RES 63 63

ACT\_SITE 183

ACT\_SITE 229

ACT\_SITE 326

FACTOR X HEAVY CHAIN.

ACTIVATED FACTOR XA, HEAVY CHAIN.

EGF-LIKE 1, CALCIUM-BINDING.

EGF-LIKE 2.

SERINE PROTEASE.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

GAMMA-CARBOXYGLUTAMIC ACID.

HYDROXYLATION (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 17 22 BY SIMILARITY.

FT DISULFID 50 61 BY SIMILARITY.

FT DISULFID 55 70 BY SIMILARITY.

FT DISULFID 72 81 BY SIMILARITY.

FT DISULFID 89 100 BY SIMILARITY.

FT DISULFID 96 109 BY SIMILARITY.

FT DISULFID 111 124 BY SIMILARITY.

FT DISULFID 132 249 INTERCHAIN (BY SIMILARITY).

FT DISULFID 148 153 BY SIMILARITY.

FT DISULFID 168 184 BY SIMILARITY.

FT DISULFID 297 311 BY SIMILARITY.

FT DISULFID 322 350 BY SIMILARITY.

FT CARBOHYD 52 52 O-LINKED.

FT CARBOHYD 186 186 N-LINKED (GLCNAC...).

SQ SEQUENCE 376 AA: 42455 MW: 75A5C7FE9BA561C1 CRC64;

Alignment Scores:

Pred. No.: 1.74e-31 Length: 376

Score: 724.00 Matches: 148

Percent Similarity: 53.25% Conservative: 73

Best Local Similarity: 35.66% Mismatches: 136

Query Match: 30.98% Indels: 58

DB: 1 Gaps: 9

US-09-763-153-1 (1-1245) x FA10\_TROCA (1-376)

QY 1 GCCAACTCTTCCTGGAGGAGCTCGTCACAGCAGCTGGAGGGAGTGCATAGAGGAG 60

Db 1 SerAsnSerLeuPheGluGluIleArgProGlyAsnIleGluArgGluCysIleGluGlu 20

QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGATGACACACTGGCCTTC 120

Db 21 LysCysSerLysGluGluAlaArgIleValPheGluAspAsnGluLysThrGluThrPhe 40

QY 121 TGGTCCAAAGCAGTCGACGCTGACCATGTGCTGTGCTTCGCTTGGACACCGCTGGCC 180

Db 41 TrpAsnValTyrValAspGlyAspGlnCysSerSerAsnPro----- 54

QY 181 AGCCTGTGCTGCGGCGCAGCGCATCGTCATCGGCGATCGGCGATCGTACGTCGCGACTGC 240

Db 55 -----CysHisTyrArgGlyThrCysLysAspGlyIleGlySerTyrThrCysThrCys 72

QY 241 CGCAGCGCTGGAGGCGCGCTTCTGCGAGCGGAGTGAGCTTCCTCAATGTCTGCTGCTG 300

Db 73 LeuProAsnTyrGluGlyLysAsnGluLys--ValLeuTyrGlnSerCysArgVal 91

QY 301 GACACGCGGCTCCACGCATTACTGCTAGAGAGAGTGGGCTGGCGCGCTGTAGCTGT 360

Db 92 AspAsnGlyAsnCysTrpHisPheCysLysArgValGlnSerGluThrGlnCysSerCys 111

QY 361 GCGCTGGCTACAAGCTGGGGGAGCAGCTCTGCTGCTGTCACCGCCAGTGAAGTTCCT 420

Db 112 AlaGluSerTyrArgLeuGlyValAspGlyHisSerCysValAlaGluCysPheSer 131

QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGGAGTCACTGCTGAAAGAGACACAGAA 480

Db 132 CysGlyArgAsnTyrLeuLysAlaArgAsnLys----- 141

QY 481 GACCAAGAAGACCAAGTAGATGCGGCTCATTTGATGGGAAGATGACACAGCGGGAGAC 540

Db 142 -----IleValAsnGlyMetAspCysLysLeuGlyGlu 152

QY 541 AGCCCTGGCAGTGTCTGCTGGACTCAAAAGAAGAGCTGGCTGGCGGGGAGCTGCTC 600

Db 153 CysProTrpGlnAlaValLeuIleAsnGluLysGlyGluValPheCysGlyClyThrIle 172

QY 601 ATCCACCCCTCTGGTGTGACAGCGGCGGCTGATGATGATGATGATGATGATGATGAT 660

Db 173 LeuSerProIleHisValLeuThrAlaHisCysIleAsnGlnThrLysSerVal--- 191

QY 661 GTCAGGCTGGAGAGTATGACCTGGCGGCTGGGAGAGTGGGAGAGTGGGAGAGTGGGAG 720

Db 192 -----LysGluThrArgArgLeu-----LeuSerVal 200



QY 721 AAGGAGGTCTTCGTCAC-----CCCAACTAC----- 747  
 Db 201 AsplyllyrValHisThrLysPheValProProAsnTyrTyrValHisGlnAsn 220  
 QY 748 ---AGCAAGAGCACCACGACATGACATGCGACTGCTGCACCTGGCCAGCCGCCACC 804  
 Db 221 PheAspArgValAlaTyrAspTyrAspPheAlaAlaIleArgMetLysThrProIleGln 240  
 QY 805 CTCTCCACACCATAGTCCCATGCTCCCGGACAGCGGCTGTCAGAGCGCGAGTC 864  
 Db 241 PheSerGluAsnValValProAlaCysLeuProThrAlaAspPheAla---AsnGluVal 259  
 QY 865 AATCAGCGCGCGCAGGAGACCTCGGACGGCTGCGGCTACACAGCAGCGCAGAGAAG 924  
 Db 260 LeuMetLysGlnAspSerGlyIleValSerGlyPheGly-----ArgIle 274  
 QY 925 GAGCGCAAGAGAACCGACCTTCCTGCTCAACTCATCAAGATTCCTGGTGTCCCGCAG 984  
 Db 275 GlnPheLysGlnProThrSerAsnThrLeuLysValIleThrValProTyrValAspArg 294  
 QY 985 AATGAGTGCAGGAGGTGATGAGCAACATGCTGTGAGAACATGCTGTGCGGCGATC 1044  
 Db 295 HisThrCysMetLeuSerSerAspPheArgIleThrGlnAsnMetPheCysAlaGlyTyr 314  
 QY 1045 CTCGGGACCGCAGGATCCCTGCGAGGCGACAGTGGGGCGCCATGTCGCTCCTTC 1104  
 Db 315 AspThrLeuProGlnAspAlaCysGlnGlyAspSerGlyProHisIleThrAlaTyr 334  
 QY 1105 CACGGCACCTGTTCTGCTGGTGGCTGTGAGTGGGTGAGGCTGTGGGCTCTTCAC 1164  
 Db 335 ArgAspThrHisPheIleThrGlyIleSerTrpGlyGluGlyCysAlaArgLysGly 354  
 QY 1165 ACTACGGCGTTTACACCAAGTACCGCTACCTACCTGAGTGCATC 1209  
 Db 355 LysTyrGlyValTyrThrLysValSerLysPheIleProTrpIle 369  
 RESULT 20  
 PRTC\_CANFA  
 ID PRTC\_CANFA STANDARD; PRT; 157 AA.  
 AC Q28278;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)  
 SE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).  
 PROC.  
 Canis familiaris (Dog).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NCBI\_TaxID=9615;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94318474; PubMed=8043441;  
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;  
 RT "A comparative study of partial primary structures of the catalytic  
 region of mammalian protein C.";  
 RL Br. J. Haematol. 86:590-600(1994).  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
 CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
 CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
 CC and VIIIA.  
 CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
 CC -!- SIMILARITY: BELONGS TO PERTIDASE FAMILY S1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D43751; BAA07808.1; -.  
 DR HSP; P04070; LPCU.  
 DR MEROPS; S01.218; -.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.  
 FT NON\_TER 1 1  
 FT ACT\_SITE 26 26 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 125 125 CHARGE RELAY SYSTEM.  
 FT DISULFID 96 110 BY SIMILARITY.  
 FT DISULFID 121 149 BY SIMILARITY.  
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON\_TER 157 157  
 SQ SEQUENCE 157 AA; 17262 MW; E8B1BACF49220DFB CRC64;  
 Alignment Scores:  
 Pred. No.: 4.19e-31 Length: 157  
 Score: 717.00 Matches: 132  
 Percent Similarity: 86.96% Conservative: 8  
 Best Local Similarity: 81.99% Mismatches: 17  
 Query Match: 30.68% Indels: 4  
 DB: 1 Gaps: 1  
 US-09-763-153-1 (1-1245) x PRTC\_CANFA (1-157)  
 QY 694 GAGAGTGGGAGCTGGACCTGGACATCAAGAGCTTCTGTCACCCCACTACAGCAAG 753  
 Db 1 GlulysGlyGluMetAspValAspIleLysGluValLeuIleHisProAsnTyrSerLys 20  
 QY 754 AGCACCCAGCAATGACATGCGCACTGTCACCTGCCAGCCGCCAGCCGCCCTCTCCGAG 813  
 Db 21 SerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaIlePheSerGln 40  
 QY 814 ACCATAGTGGCCATCTCCCTCCCGACAGCGGCTTTCAGAGCGGAGCTCAATCAGGCC 873  
 Db 41 ThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuThrGlnVal 60  
 QY 874 GCGCAGGAGACCTCTGACGGGCTGGGCTACACAGACCGCAGAGAGGAGGCGCAAG 933  
 Db 61 GlyGlnGluThrValValThrGlyTyrGlyTyrArgSer-----GluThrLys 76  
 QY 934 AGAACCCGACCTCTGCTCAACTTCATCAAGATTCCTGTCGCCGACCAATGAGTGC 993  
 Db 77 ArgAsnArgThrPheValLeuAsnPheIleAsnIleProValAlaProHisAsnGluCys 96  
 QY 994 AGCAGGTGATGAGCAACATGTTGTGTGAGAATGCTGTGTGGGGCATCTCTCGGGAC 1053  
 Db 97 IleGlnAlaMetTyrAsnMetIleSerGluAsnMetLeuCysAlaGlyIleLeuGlyAsp 116  
 QY 1054 CGCAGGATGCGGCGGAGCGGACAGTGGGGGCCCATGTCGCTCTTCACAGGAC 1113  
 Db 117 SerArgAspAlaCysGluGlyAspSerGlyTyrGlyProMetValThrSerPheArgGlyThr 136  
 QY 1114 TGGTTCCTGTTGGGCTGTTGAGTGGGCTGTGGGCTGTGGCTCTTCACTACAGTGGC 1173  
 Db 137 TrpPheLeuValGlyLeuValSerTrpGlyGlyGlyCysGlyArgLeuHisAsnTyrGly 156  
 QY 1174 GTT 1176  
 Db 157 Ile 157

Search completed: March 17, 2003, 17:32:35  
 Job time : 57 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:24:24 ; Search time 73.5 Seconds  
(without alignments)  
6980.372 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccactctcttcggagga.....acaaggaagccccccagaag 1245

Bring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool/US09763153/runat\_11032003\_084247\_1993/app\_query.fasta.1.1415  
-DB-SPREMBL\_21 -QMT-fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=60 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153@cgn.1.1.86@runat\_11032003\_084247\_1993 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_Organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1890.5	80.9	456	6	Q9TTRO	Q9TTRO canis famil

2	1631	69.8	460	11	Q91WN8	Q91wn8 mus musculus
3	1625	69.5	460	11	Q99PC6	Q99pc6 mus musculus
4	818.5	35.0	482	11	Q63207	Q63207 rattus norv
5	794	34.0	481	11	Q88947	Q88947 mus musculus
6	793	33.9	481	11	O54740	O54740 mus musculus
7	793	33.9	481	11	Q99L32	Q99l32 mus musculus
8	787	33.7	433	13	Q90YK1	Q90yK1 brachydanio
9	781	33.4	701	4	Q96PQ8	Q96pQ8 homo sapien
10	773	33.1	469	6	Q9GMD9	Q9gmd9 ornithorhyn
11	772	33.0	446	11	Q61109	Q61109 mus musculus
12	740	31.7	461	6	Q95ND7	Q95nd7 pan troglod
13	736	31.5	456	4	Q14316	Q14316 homo sapien
14	718	30.7	461	6	Q95ND6	Q95nd6 pan troglod
15	525	22.5	608	13	Q9PTW7	Q9ptW7 struthio ca
16	512	21.9	607	13	Q91001	Q91001 gallus gall
17	510.5	21.8	799	11	Q9DB10	Q9db10 mus musculus
18	477.5	20.4	653	11	Q8VCS4	Q8vcs4 mus musculus
19	470	20.1	399	11	Q9CQW3	Q9cqW3 mus musculus
20	440.5	18.8	1379	5	Q9V4N6	Q9v4N6 drosophila
21	433	18.5	624	11	Q9DAT3	Q9dat3 mus musculus
22	431.5	18.5	685	11	Q91WP0	Q91wp0 mus musculus
23	431	18.4	387	5	Q9XV57	Q9xy57 ctenocephal
24	427.5	18.3	420	13	Q90504	Q90504 eptatretus
25	427	18.3	624	11	Q91Y47	Q91y47 mus musculus
26	426	18.2	685	11	Q92338	Q92338 mus musculus
27	424.5	18.2	701	11	Q9JUS9	Q9jjs9 rattus norv
28	422	18.1	767	13	Q9DGR2	Q9dgr2 xenopus lae
29	420	18.0	638	11	Q8ROP5	Q8rop5 mus musculus
30	420	18.0	678	11	Q9JUS8	Q9jjs8 rattus norv
31	419.5	18.0	277	5	Q96899	Q96899 scolopendra
32	418.5	17.9	855	11	Q9JJI7	Q9jji7 rattus norv
33	417.5	17.9	643	6	Q97306	Q97306 sus scrofa
34	417.5	17.9	812	11	Q91WJ5	Q91wj5 mus musculus
35	417	17.8	573	5	Q9V516	Q9v516 drosophila
36	416	17.8	1186	5	Q9VSU2	Q9vsu2 drosophila
37	416	17.8	1449	5	Q9UII2	Q9uli2 drosophila
38	416	17.8	1462	5	Q9UII3	Q9uli3 drosophila
39	416	17.8	2382	5	Q9B1I9	Q9bli9 drosophila
40	416	17.8	2409	5	Q960G6	Q960g6 drosophila
41	415.5	17.8	453	11	Q8VDE0	Q8vde0 mus musculus
42	415	17.8	541	11	Q9QX90	Q9qx90 rattus norv
43	415	17.8	623	11	Q9JJP3	Q9jjp3 rattus norv
44	415	17.8	643	11	Q9QX84	Q9qx84 rattus norv
45	414	17.7	435	5	Q9NEY2	Q9nfy2 anopheles g
46	413.5	17.7	366	11	Q9QX85	Q9qx85 rattus norv
47	413	17.7	490	11	Q920K3	Q920k3 rattus norv
48	411	17.6	810	4	Q15146	Q15146 homo sapien
49	410	17.5	560	4	Q14520	Q14520 homo sapien
50	410	17.5	787	5	Q9VEV6	Q9vev6 drosophila
51	409	17.5	467	5	Q967X8	Q967x8 panulirus a
52	409	17.5	698	13	Q9P071	Q9pu71 xenopus lae
53	408	17.5	722	13	Q9PS25	Q9ps25 lampetra ja
54	407	17.4	334	6	Q46507	Q46507 papio hamad
55	406.5	17.4	275	4	Q96RZ6	Q96rz6 homo sapien
56	406.5	17.4	1047	5	Q9VZH2	Q9vzh2 drosophila
57	404.5	17.3	581	5	Q960I5	Q960i5 drosophila
58	404.5	17.3	733	5	Q9VYX9	Q9vtx9 drosophila
59	404.5	17.3	1185	5	Q9VW38	Q9vw38 drosophila
60	404	17.3	246	11	Q9Z1R9	Q9z1r9 mus musculus

## ALIGNMENTS

RESULT 1  
Q9TTRO ID Q9TTRO PRELIMINARY; PR; 456 AA.  
AC Q9TTRO;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)  
GN Protein C precursor.  
OS Canis familiaris (Dog).



## RESULT 2

Q91WN8 PRELIMINARY; PRT; 460 AA.  
 AC Q91WN8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to protein C.  
 GN PROC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 SEQUENCE FROM N.A.  
 TISSUE=LIVER;  
 Strausberg R.;  
 Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013896; AAH13896.1; -.  
 DR MGD; MGI:97771; Proc.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_2.  
 DR PROSITE; PS01187; EGF\_CA; UNKNOWN\_1.  
 DR PROSITE; PS00011; GLU CARBOXYLATION; UNKNOWN\_1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 460 AA; 51818 MW; 0117F26E68FCC274 CRC64;

## Alignment Scores:

Pred. No.: 1.55e-110 Length: 460  
 Score: 1631.00 Matches: 290  
 Percent Similarity: 83.74% Conservative: 55  
 Best Local Similarity: 70.39% Mismatches: 63  
 Query Match: 69.79% Indels: 4  
 DB: 11 Gaps: 3

us-09-763-153-1 (1-1245) x Q91WN8 (1-460)

QY 1 GCCAACTCTCTCTGAGGAGCTCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 DB 42 AlaAsnSerPheLeuGluGluMetArgProGlySerLeuGluArgGluCysMetGluGlu 61  
 QY 61 ATCTGTGACTTCGAGAGCGCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 DB 62 IleCysAspPheGluGluAlaGlnGluPheGlnAsnValGluAspThrLeuAlaPhe 81  
 QY 121 TGGTCCAAAGCAGCTCGAGCGTGACAGGTGTTGTTGCTTGGCCCTGGAGCAGCCGTCGCC 180  
 DB 82 TrpIleLysTyrPheAspGlyAspGlnCysSerAlaProProLeuAspHisGlnCysAsp 101  
 QY 181 AGCTGTGCTCGGGCAGCGCAGCTGTCATCGACGCATCGGCAGCTTCAGCTGCAGCTGC 240  
 DB 102 SerProCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysSerCys 121  
 QY 241 CGCAGCGCTGGAGGGCGCTTCCTGACGCGCGAGGTGAGCTTCCTCAATGTGCTGCTG 300  
 DB 122 AspLysGlyTrpGluGlyLysPheCysGlnGlnGluLeuArgPheGlnAspCysArgVal 141  
 QY 301 GACACAGCGGCTGTCAGCATTACTGCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360  
 DB 142 AsnAsnGlyGlyCysLeuHisTyrCysLeuGluGluSerAsnGlyArgArgCysAlaCys 161

QY 361 CGCCTGCTTACAAAGCTGGGGGAGCAGCCTCTCTGAGTGTACCCCGGAGTGAAGTCCCT 420  
 DB 162 AlaProGlyTyrGluLeuAlaAspHisMetArgCysLysSerThrValAsnPhePro 181  
 QY 421 TGTGGAGGCGCTGGAAGCGGATGGAAGAAGCGCAGTCACTCACTGAAACGAGACACAGAA 480  
 DB 182 CysGlyLysLeuGlyArgTrpIleGluLysArgLysIleLeuLysArgAspThr--- 200  
 QY 481 GACCAAGAAGACCAAGTA-----GATCCGCGCTCATTTGATGGGAAGATGACCAAGCGG 534  
 DB 201 AspLeuGluAspGluLeuGluProAspProArgIleValAsnGlyThrLeuThrLysGln 220  
 QY 535 GGACACAGCCCTGGCAGGTGGTCTCTGCTGGACTCAAGAAGAAGCTGGCTGGCGGGGCA 594  
 DB 221 GlyAspSerProTrpGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGlyGly 240  
 QY 595 GTGCTCATCCACCCTCTCTGGTGTCTGACAGCGGCCACCTGCATGGATGATGATCAACAAG 654  
 DB 241 ValLeuIleHisThrSerTrpValLeuThrAlaAlaHisCysValGluGlyThrLysLys 260  
 QY 655 CTCCTTGTGAGCTTGGAGAGTATGACCTGGCGGCTGGAGAGTGGAGCTGGACCTG 714  
 DB 261 LeuThrValArgLeuGlyGluTyrAspLeuArgArgAspHisTrpGluLeuAspLeu 280  
 QY 715 GACATCAAGGAGGCTTCTGCCACCCCACTACAGCAGAGCAGCACCACCGACATGACATC 774  
 DB 281 AspIleLysGluIleLeuValHisProAspTyrThrArgSerSerAspAsnAspIle 300  
 QY 775 GCATGCTGCACCTTGGCCCGCCAGCCCTCTTCGCAGACCATAGTGCCTTCCTGCTC 834  
 DB 301 AlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 320  
 QY 835 CCGACAGCGGCTTGGAGCGGAGCTCAATCAGCGCGCCGAGAGACCCCTCGTGACG 894  
 DB 321 ProAsnAsnGlyLeuAlaGln---GluLeuThrGlnAlaGlyGlnGluThrValValThr 339  
 QY 895 GGCTGGGCTACCCAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954  
 DB 340 GlyTrpGlyTyrGlnSerAspArgIleLysAspGlyArgArgAsnArgThrPheIleLeu 359  
 QY 955 AACTTCATCAAGATTCCTGCTCCCGCCAGCAATGAGTGCAGCGAGGTGCATGAGCAACATG 1014  
 DB 360 ThrPheIleArgIleProLeuValAlaArgAsnGluCysValGluValMetLysAsnVal 379  
 QY 1015 GTGTCTGAGACATGCTGTGTGGGCGATCCTCGGGAGCCGCGAGGAGTGTCTGGAGGC 1074  
 DB 380 ValSerGluAsnMetLeuCysAlaGlyIleIleGlyAspThrArgAspAlaCysAspGly 399  
 QY 1075 GACAGTGGGGCGCCATGCTGCTCTCCACGCGACCTGCTCTGTTGGTGGCGCTGCTG 1134  
 DB 400 AspSerGlyGlyProMetValValPhePheArgGlyThrTrpPheLeuValGlyLeuVal 419  
 QY 1135 AGCTGGGGTGGAGGCTGTGGGCTCTTCACTACAGTACGGCGTGTACACCAAGTACGCGCG 1194  
 DB 420 SerTrpGlyGluGlyCysGlyHisThrAsnAsnTyrGlyIleTyrThrLysValGlySer 439  
 QY 1195 TACCTCCACTGGATCCATGGCAGCATCAGAGACAAG 1230  
 DB 440 TyrLeuLysTrpIleHisSerTyrIleGlyGluLys 451  
 RESULT 3  
 Q99PC6 PRELIMINARY; PRT; 460 AA.  
 ID Q99PC6;  
 AC Q99PC6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Anticoagulant protein C.  
 GN PROC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

```

RN 1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Korf 1.;
RT "Complete sequence of UC72A01."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF318182; AAK07918.1; -.
DR HSP; P04070; IPCU.
DR MEROPS; S01.218; -.
DR MGD; MGI:97771; Proc.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_Blood.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00001; EGF_Like; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_Ca; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW Hydroxylation; Repeat; Serine protease.
SQ SEQUENCE 460 AA; 51784 MW; 0293BC25E9D3ED16 CRC64;

```

```

Alignment Scores:
Pred. No.: 4, 24e-110 Length: 460
Score: 1625.00 Matches: 289
Percent Similarity: 83.50% Conservative: 55
Best Local Similarity: 70.15% Mismatches: 64
Query Match: 69.53% Indels: 4
DB: Gaps: 3

```

US-09-763-153-1 (1-1245) x Q99PC6 (1-460)

```

QY 1 GCCAACTCCTCTGGAGAGCTCCGTCACAGCCCTGGAGCGGAGTGCATAGAGAG 60
|||||
42 AlaAsnSerPheLeuGluGluMetArgProGlySerLeuGluArgGluCysMetGluGlu 61
61 ATCTGTGACTTCAGAGCGCCAGCAAAATTTCCAAATGCGATGACACACTGGCCTTC 120
|||||
62 IleCysAspLeuGluGluAlaGlnGluIlePheGlnAsnValGluAspThrLeuAlaPhe 81
|||||
QY 121 TGGTCCAGCAGCTCGAGCGGTGACCACTGCTGGTCTTGGCCCTGGAGCACCCGTCGCC 180
|||||
DB 82 TrpIleLysTyrPheAspGlyAspGlnCysSerAlaProLeuAspHisGlnCysAsp 101
QY 181 AGCCGTGTGTCGGCGCAGCGGACGTCATGTCAGCGCATCGGCAGCTTCAGCTGGCAGTC 240
|||||
DB 102 SerProCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysSerCys 121
QY 241 CGCAGCGCTGGAGCGCGCTCTGCGCAGCGGAGGTGAGCTTCCTCAATGTGCTCGCTG 300
|||||
DB 122 AspLysGlyTrpGlnGlyLysPheCysGlnGlnGluLeuArgPheGlnAspCysArgVal 141
QY 301 GACAAGCGCGGTGACGCAATTCCTAGAGAGGTGGGTGGCGCGCTGTAGCTGT 360
|||||
DB 142 AsnAsnGlyGlyCysLeuHisTyrCysLeuGluGluSerAsnGlyArgArgCysAlaCys 161
|||||

```

```

QY 361 GGCCTGCTCAAGCTGGGAGCAGCAGCTCCTGCACTGTCACCCGCGAGTGAAGTTCCT 420
|||||
DB 162 AlaProGlyTyrGluLeuAlaAspHisMetArgCysLysSerThrValAsnPhePro 181
|||||
QY 421 TGTGGAGCCCTGGAGCGGATGAGAGAGCGCAGTCACTCACTGAACGAGACACAGAA 480
|||||
DB 182 CysGlyLysLeuGlyArgTrpIleGluLysLysArgLysIleLeuLysArgAspThr 200
|||||
QY 481 GACCAGAGACCAAGTA-----GATCCGCGCTCATTTGATGGGAAGATGACCGGGG 534
|||||
DB 201 AspLeuGluAspGluLeuGluProAspProArgIleValAsnGlyThrLeuThrLysGln 220
|||||
QY 535 GGAGACAGCCCTGGCAGCGTCTCTGCTGCTGAGTCACTCAAGAGAAAGTGGCTCGGGGCA 594
|||||
DB 221 GlyAspSerProTrpGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGlyGly 240
|||||
QY 595 GTGCTCATCCACCCCTCTGGGTGCTGAGCAGCGGCCCTGCTGATGATGATGATGATGATG 654
|||||
DB 241 ValLeuIleHisThrTrpValLeuThrAlaAlaHisCysValGluGlyThrLysLys 260
|||||
QY 655 CTCCTTGTGAGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGGAGTGGACCTG 714
|||||
DB 261 LeuThrValArgLeuGlyGluTyrAspLeuArgArgAspHisTrpIleLeuAspLeu 280
|||||
QY 715 GACATCAAGGAGTCTCTCTCCACCCCACTACAGCAAGAGCACCACCGCAATGACATC 774
|||||
DB 281 AspIleLysGluIleLeuValHisProAsnTyrThrArgSerSerSerAspAsnAsp 300
|||||
QY 775 GCACTGCTGACCTGGCCCGCCCGCCCTCTCCACACCATAGTGGCCCTCTGCTC 834
|||||
DB 301 AlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 320
|||||
QY 835 CCGACAGCGGCTTCGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTCTGCTGAG 894
|||||
DB 321 ProAsnGlyLeuAlaGln---GluLeuThrGlnAlaGlyGlnGluThrValThr 339
|||||
QY 895 GGTGCGGTACACAGCAGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954
|||||
DB 340 GlyTrpGlyTyrGlnSerAspArgIleLysAspGlyArgArgAsnArgThrPheIleLeu 359
|||||
QY 955 AACTTCATCAAGATCCCGTGTCCCGCACAATGAGTGCAGCGAGCTCATGACCAATG 1014
|||||
DB 360 ThrPheIleArgIleProLeuValAlaArgAsnGluCysValGluValMetLysAsnVal 379
|||||
QY 1015 GTCTCTGAGAACATGTGTGCGGGCATCTCGGGGACCGGCGGAGGATCGCTCGAGGCG 1074
|||||
DB 380 ValSerGluAsnMetLeuCysAlaGlyIleIleGlyAspThrArgAspAlaCysAspGly 399
|||||
QY 1075 GACAGTGGGGGGCGGCTGCTCCTCTCCACGCGCAGCTGCTCTGCTGGGCGCTGGTG 1134
|||||
DB 400 AspSerGlyGlyProMetValValPhePheArgGlyThrTrpPheLeuValGlyLeuVal 419
|||||
QY 1135 AGCTGGGTGAGGCTGTGGGCTCTCTCAACACTAGCGGCTTTTACACCAAGTACCGCCG 1194
|||||
DB 420 SerTrpGlyGlyCysGlyHisThrAsnAsnTyrGlyIleThrThrLysValGlySer 439
|||||
QY 1195 TACCTGACGTGGATCCATGGGACATCAGAGACAAG 1230
|||||
DB 440 TyrLeuLysTrpIleHisSerTyrIleGlyGluLys 451
|||||
RESULT 4
Q63207 PRELIMINARY; PRT; 482 AA.
ID AC Q63207;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor X.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

```





ID	Accession	PRELIMINARY	PRT	481 AA
AD	088947			
AC	088947			
DT	01-NOV-1998	(TREMBLrel. 08, Created)		
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)		
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)		
DE		Coagulation factor X precursor.		
DE		F10.		
GN		GN		
OS		Mus musculus (Mouse).		
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX		NCBI_TaxID=10090;		
RN		[1]		
RC		SEQUENCE FROM N.A.		
RP		STRAIN=C57BL/6 X CBA; TISSUE=LIVER;		
RX		MEDLINE=98347933; PubMed=9684791;		
RA		Liang Z., Cooper A., DeFord M.E., Carmeliet P., Collen D.,		
RA		Castellino F.J., Rosen E.D.;		
RT		"Cloning and characterization of a cDNA encoding murine coagulation		
RT		factor X.;"		
RT		Thromb. Haemost. 80:87-91(1998).		
RY		[2]		
SC		SEQUENCE FROM N.A.		
RC		STRAIN=129Sv;		
RA		Cooper A., Liang Z., Castellino F.J., Rosen E.D.;		
RT		"Cloning and Characterization of the Murine Factor X Gene.;"		
RL		Thromb. Haemost. 0:0-0(2000).		
CC		-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE		
CC		TRYPSIN FAMILY.		
EMBL		AF087644; AAC36345.1; -.		
DR		EMBL; AF211347; AAF22980.1; -.		
DR		HSSP; P00742; 1XKA.		
DR		MEROPS; S01.216; -.		
DR		MGD; MGI:103107; F10.		
DR		InterPro; IPR000152; Asx_hydroxyl.		
DR		InterPro; IPR001314; Chymotrypsin.		
DR		InterPro; IPR000561; EGF-like.		
DR		InterPro; IPR000742; EGF_2.		
DR		InterPro; IPR001881; EGF_Ca.		
DR		InterPro; IPR002383; GLA_blood.		
DR		InterPro; IPR001254; Ser_protease_Try.		
DR		InterPro; IPR000294; Vitk_dep_GLA.		
DR		Pfam; PF00008; EGF; 2.		
DR		Pfam; PF00594; gla; 1.		
DR		Pfam; PF00089; trypsin; 1.		
DR		PRINTS; P00722; CHYMOTRYPSIN.		
DR		PRINTS; PR00001; GLABLOOD.		
DR		SMART; SM00179; EGF_CA; 1.		
DR		SMART; SM00001; EGF_like; 1.		
DR		SMART; SM00069; GLA; 1.		
DR		SMART; SM00020; Tryp_Spc; 1.		
DR		PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.		
DR		PROSITE; PS00022; EGF_1; UNKNOWN_1.		
DR		PROSITE; PS01186; EGF_2; 2.		
DR		PROSITE; PS01187; EGF_CA; 1.		
DR		PROSITE; PS00011; GLU CARBOXYLATION; 1.		
DR		PROSITE; PS50240; TRYPSIN_DOM; 1.		
DR		PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.		
DR		PROSITE; PS00135; TRYPSIN_SER; 1.		
KW		Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;		
FT		Serine protease; Signal.		
FT		SIGNAL 1 40		
FT		CHAIN 41 481		
FT		COAGULATION FACTOR X.		
SQ		SEQUENCE 481 AA; 54018 MW; 8AC09DE5F9D271E CRC64;		
Alignment Scores:				
Pred. No.:		1,31e-49	Length:	481
Score:		794.00	Matches:	162
Percent Similarity:		53.71%	Conservative:	77
Best Local Similarity:		36.40%	Mismatches:	158
Query Match:		33.98%	Indels:	48
DB:		11	Gaps:	9

QY 985 AATGAGTGCAGGAGTCTATGACAAACATGGTCTGTGAGAAACATGCTGTGGGGCATC 1044  
 DB 385 AsnThrCysLysLeuSerThrSerPheSerIleThrGlnAsnMetPheCysAlaGlyTyr 404  
 QY 1045 CTGGGGACCGGAGAGTCTCGAGGGCGAGAGTGGGGGCCCGCCATGTCGCTCTTC 1104  
 DB 405 GluAlaLysLeuGluAspAlaCysGlnGlyAspSerGlyGlyProHisValThrArgPhe 424  
 QY 1105 CACGGACCTGGTCTGTGGCCCTGGTGGAGTGGGGTGGAGGCTGGTGGCTCTCTAC 1164  
 DB 425 LysAsnThrTyrThrValThrGlyIleValSerTrpGlyGlyCysAlaArgLysGly 444  
 QY 1165 AACTACGGCGTTTACACCAAGTACCGCTCTCGACTGGATCCATGGGCACATCAGA 1224  
 DB 445 LysTyrGlyIleThrLysValThrThrPheLeuLysTrpIleAspArgSerMetLys 464  
 QY 1225 GACAAGGAGCCGCC 1239  
 DB 465 AlaArgValGlyPro 469

RESULT 6  
 O54740 PRELIMINARY; PRT; 481 AA.  
 AC O54740;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Coagulation factor X precursor (EC 3.4.21.6).  
 GN F10 OR FAL0.  
 OS Mus musculus (Mouse).  
 OG Plasmid pBluescript.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE=98454993; PubMed=9783672;  
 RA Heidtmann H.H., Kontermann R.E.;  
 RT "Cloning and recombinant expression of mouse coagulation factor X";  
 RL Thromb. Res. 92:33-41(1998).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
 DR EMBL; AJ222677; CAA10933.1; -.  
 DR HSSP; P00742; 1XRA.  
 DR MEROPS; S01.216; -.  
 DR MGD; MGI:103107; F10.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00072; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF\_like; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Plasmid;  
 KW Repeat; Serine protease; Signal.  
 FT SIGNAL 1 40 POTENTIAL.  
 FT CHAIN 41 481 COAGULATION FACTOR X.  
 SQ SEQUENCE 481 AA; 53986 MW; CF702DESEF9D97AE CRC64;  
 Alignment Scores:  
 Pred. No.: 1.54e-49 Length: 481  
 Score: 793.00 Matches: 162  
 Percent Similarity: 53.71% Conservative: 77  
 Best local Similarity: 36.40% Mismatches: 158  
 Query Match: 33.93% Indels: 48  
 DB: 11 Gaps: 9  
 US-09-763-153-1 (1-1245) x O54740 (1-481)  
 QY 1 GCCAACTCTCTGGAGGAGTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
 DB 41 AlaAsnSerPhePheGluGluPheLysLysGlyAsnLeuGluArgGluCysMetGluGlu 60  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 DB 61 IleCysSerTyrGluGluValArgGluIlePheGluAspAspGluLysThrLysGluTyr 80  
 QY 121 TGGTCCAAAGCAGTCGAGCGTGACAGTGTGGTCTTGCCCTTGGAGCACCCCTGCGCC 180  
 DB 81 TrpThrLysTyrLysAspGlyAspGlnCysGluSerSerPro----- 94  
 QY 181 AGCCTGTGCTCGGGCGACGCGACGTGCATCGCGCATCGGACGTTCAGTCGCGACTG 240  
 DB 95 -----CysGlnAsnGlnGlyAlaCysArgAspGlyIleGlyTyrThrCysThrCys 112  
 QY 241 CGCAGCGCTGGGAGGCGCGCTTCTGCAGCGCGAGGTGAGCTTCCTCAATGTCGCTG 300  
 DB 113 SerGluGlyPheGluGlyLysAsnCysGluLeuPheValArgLysLeu--CysArgLeu 131  
 QY 301 GACAACGGCGCTGCACGCATTACTGCTAGAGGAGTGGGCTGGCGCGCTGTAGCTGT 360  
 DB 132 AspAsnGlyAspCysAspGlnPheCysArgGluGluGlnAsnSerValValCysSerCys 151  
 QY 361 CGGCTGTGCTACAAGCTGGGGAGGACCTCTCGAGTGTACCCCGCAGTGAAGTCCCT 420  
 DB 152 AlaSerGlyTyrPheLeuGlyAsnAspGlyLysSerCysIleSerThrAlaProPhePro 171  
 QY 421 TGTGGAGGCCCTCGAAGCGGATGGAGAGAG-----CGCAGTCACTG 465  
 DB 172 CysGlyLysIleThrThrGlyArgArgLysArgSerValAlaLeuAsnThrSerAspSer 191  
 QY 466 AAACGACACACAGAGAC-----CAAGAAGACCAAGTAGATCCG----- 504  
 DB 192 GluLeuAspLeuGluAspAlaLeuLeuAspGluAspPheLeuSerProThrGluAsnPro 211  
 QY 505 -----CGG 507  
 DB 212 IleGluLeuLeuAsnLeuAsnGluThrGlnProGluArgSerSerAspAspLeuValArg 231  
 QY 508 CTCATTGATGGAGATGACACCGCGGGAGACAGCCCTGGAGGTGGTCTGCTCGAC 567  
 DB 232 IleValGlyGlyArgGluCysLysAspGlyCysProTrpGlnAlaLeuLeuIleAsn 251  
 QY 568 TCAAAGAGAGAGCTGGCTGGCGGAGTGTCTATCCACCCCTCTCGGTGGTGCACAGCG 627  
 DB 252 GluAspAsnGluGlyPheCysGlyThrIleLeuAsnGluPheTyrIleLeuThrAla 271  
 QY 628 GCCACTGCTGATGATGATGATCCAGAGCTCTCTGTAGGCTTGAGAGTATGACCTGCGG 687  
 DB 272 AlaHisCysLeuHisGlnAlaArgArgPheLysValArgValGlyAspArgAsnThrGlu 291  
 QY 688 CGCTGGGAGAAGTGGGAGCTGGACCTGGACATCAAGAGAGCTTCGTCACACCCCACTAC 747  
 DB 292 LysGluAspGlyAsnGluLeuValHisGluValAspValIleLysHisAsnLysPhe 311  
 QY 748 AGCAAGAGCACCACCGCAATGACATCGCACTGTGTCACCTGGCGCCACCGCCGCCCTC 807





Db 259 GluGlyThrGluGlnLeuLeuGlnValAspGlnMetPheThrHisProAlaTyrValSer 278  
QY 754 AGCACCCAGCAATAGATCGACTGCTGCACCTGGCCAGCCCGCCAGCCCTCTCGAG 813  
Db 279 GluThrAlaAspSerAspAlaLeuLeuArgLeuThrProIleValTyrSerVal 298  
QY 814 ACCATAGTCCCATCTGCTCCCGACAGCCGCTTGCAGAGCCGAGCTCAATCAGGCC 873  
Db 299 TyrAlaValProValCysLeuProLeuArgGluMetAlaGluArgGluLeuTrpAlaVal 318  
QY 874 GGCCAGGAGACCTCGTCCAGCGGTGGGCTACACAGCAGCCGAGAGAGGCCCAAG 933  
Db 319 SerLysHisThr---ValSerGlyTrpGly-----LysArgSerGluAspGlyPro 334  
QY 934 AGAAACCCGACCTTGGTCTCTCAATCATCAAGATTCCGCTGGTCCGCGACAAATGAGTGC 993  
Db 335 ThrSerArg-----LeuLeuArgArgLeuLeuValProArgIleArgThrGlnGluCys 352  
QY 994 AGCGAGTTCATGAGCAACATGGTGTCTGAGAACATGTGTGTGGGCGATCTCTGGGGAC 1053  
Db 353 ValGlnValSerAsnLeuThrLeuThrSerAsnMetPheCysAlaGlyTyrIleGluGly 372  
1054 CGCGAGGATGCTCGAGGCGGAGAGTGGGGCCGATGGTGTCTCTCCAGCGGCACC 1113  
Db 373 ArgGlnAspSerCysLysGlyAspSerGlyGlyProLeuValThrArgTyrArgAspThr 392  
QY 1114 TGGTTCCTGGTGGGCTGCTGAGTGGGGTGGGCTGGTGGCTCTCAACATACAGGC 1173  
Db 393 AlaPheLeuLeuGlyIleValSerTrpGlyLysGlyCysAlaArgProGlySerTyrGly 412  
QY 1174 GTTTACACAAAGTCAGCGCTACCTCGACTGGATC 1209  
Db 413 IleTyrThrArgValSerAsnTyrLeuGlnTrpIle 424  
RESULT 9  
Q96PQ8  
ID Q96PQ8 PRELIMINARY; PRT; 701 AA.  
AC Q96PQ8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Factor VII active site mutant immunconjugate.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. PubMed=11593034;  
RX MEDLINE=21477446; PubMed=11593034;  
RA Hu Z., Garen A.;  
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor  
cells for immunotherapy in mouse models of prostatic cancer."  
Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).  
EMBL: AF272774; AAK58686.1; -  
InterPro; IPR000152; Asx\_hydroxyl.  
InterPro; IPR000561; EGF-like.  
InterPro; IPR000742; EGF\_2.  
InterPro; IPR001881; EGF\_Ca.  
InterPro; IPR003006; Ig\_MHC.  
InterPro; IPR001254; Ser\_protease\_Try.  
InterPro; IPR000294; VitK\_dep\_GLA.  
Pfam; PF00008; EGF; 2.  
Pfam; PF00594; gla; 1.  
Pfam; PF00047; Ig; 2.  
Pfam; PF00089; trypsin; 1.  
SMART; SM00181; EGF; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
DR PROSITE; PS01187; EGF\_Ca; UNKNOWN\_1.  
DR PROSITE; PS00011; GLU\_CARBOXYLATION; UNKNOWN\_1.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
KW PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;  
Alignment Scores:  
Pred. No.: 1.23e-48 Length: 701  
Score: 781.00 Matches: 164  
Percent Similarity: 56.50% Conservative: 75  
Best Local Similarity: 38.77% Mismatches: 148  
Query Match: 33.42% Indels: 36  
DB: 4 Gaps: 10  
US-09-763-153-1 (1-1245) x Q96PQ8 (1-701)  
QY 1 GCCAATCTCTCTCTGAGGAGTCCGCTCAGCAGCAGCTGGAGCGGAGTGCATAGAGAG 60  
Db 61 AlaAsnAlaPheLeuGluLeuArgProGlySerLeuGluArgGluCysLysGluGlu 80  
QY 61 ATCTGTCACTTCGAGGAGGCCCAAGAAATTTCCAAATATGTGATGACACACTGGCCTTC 120  
Db 81 GlnCysSerPheGluGluAlaArgGluIlePheLysAspAlaGluArgThrLysLeuPhe 100  
QY 121 TGTCTCAAGACGCTCAGCGTGACAGTGTCTGCTTCCCTTGCAGGACCCGTCGCGC 180  
Db 101 TrpIleSerTyrSerAspGlyAspGlnCys-----AlaSer 112  
QY 181 AGCTGTGTGGCGGACGCGCAGTGCATCGAGCGCATCGGAGCTTCAGTCCGACTGC 240  
Db 113 SerProCysGlnAsnGlyLysSerCysLysAspGlnLeuGlnSerTyrIleCysPheCys 132  
QY 241 CGCAGCGGTGGAGGGCGGCTTCTGCCAG---CGCAGAGGTGAGCTTCCTCAATGCTCG 297  
Db 133 LeuProAlaPheGluGlyArgAsnCysGluThrHisLysAspAspGlnLeuIleCysVal 152  
QY 298 CTGACAAAGCGGCTGACGCAATTAATCTACTGCTAGAGAGAGTGGGCTGGCGGCGC ---TGT 354  
Db 153 AsnGluAsnGlyGlyCysGluGlnTyrCysSerAspHisThrGlyThrLysArgSerCys 172  
QY 355 AGCTGTGGCGCTGCTACAGCTGAGGCGGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 414  
Db 173 ArgCysHisGluGlyTyrSerLeuLeuAlaAspGlyValSerCysThrProThrValGlu 192  
QY 415 TTCCTTGTGGAGCGCTGGAAAGCGGATGGAGAAAGCGGAGTGCATGCTGAAAGAGAC 474  
Db 193 TyrProCysGlyLys---IleProIleLeuGluLysArgAsnAla----- 206  
QY 475 ACAGAGACCAAGAACAGCAAGTAGATCGCGCTCATTTGATGGGAGAGTACCAGCGCG 534  
Db 207 -----SerLysProGlnGlyArgIleValGlyGlyLysValCysProLys 221  
QY 535 GGAGACAGCCCTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594  
Db 222 GlyGlyCysProThrGlnValLeuLeuValAsnGlnValGlnLeu---CysGlyGly 240  
QY 595 GTGCTCATCACCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651  
Db 241 ThrLeuIleAsnThrIleTrpValValSerAlaAlaHisCysPheAspLysLysAsn 260  
QY 652 -----AAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705  
Db 261 TrpArgAsnLeuIleAlaValLeuGlyGluHisAspLeuSerGluHisAspGlyAspGlu 280  
QY 706 CTGGAGCTGGACATCAAGGAGGTCTTCTGCTCCACCCCACTACAGCAAGAGCACCAGCCGAC 765  
Db 281 GlnSerArgValAlaGlnValIleIleProSerThrTyrValProGlyThrThrAsn 300  
QY 766 AATGACATCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825  
Db 301 HisAspIleAlaLeuLeuArgLeuHisGlnProValValLeuThrAspHisValValPro 320  
QY 826 ATCTGCTCCCGGACAGCGGCTTGGCAGAGCGGAGCTCAATCAGGCGCGGAGGAGAC 885  
:::|||||::: |||



Db 321 LeuCysLeuProGluArgThrPheSerGluArgThrLeuAlaPheVal----ArgPheSer 339  
 QY 886 CTCGTGACGGCTGGGGCTACACAGCAGCCGAGAGAGAGGCCAAGAAACCGCAC 945  
 |||||:|||||  
 Db 340 LeuValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla----- 354  
 QY 946 TTCGTCTCAACTTCATCAAGATTCCCGTGGTCCCGCACAATGAGTGCAGCGAGTCA 1005  
 |||||:|||||  
 Db 355 LeuGluLeuMetValLeuAsnValProArgLeuMetThrGlnAspCysLeuGlnGlnSer 374  
 QY 1006 -----AGCAACATGGTCTGTGAGAACATGCTGTCTCGGGCATCTCCGG 1050  
 |||||:|||||  
 Db 375 ArgLysValcylAspSerProAsnIleThrGluTyrMetPheCysAlaGlyTyrSerAsp 394  
 1051 GACCGCAGATGCTCGCAGGCGACAGTGGGGGCCCATGGTCCGCTCTCCACGCGC 1110  
 :|||||:|||||  
 395 GlySerLysAspSerCysAlaGlyAspSerGlyGlyProHisAlaThrHisTyrArgGly 414  
 QY 1111 ACTGTGCTTCTGTGGGCTGTGAGTGGGTGAGGGTGTGGGCTCTCTCCACAACTAC 1170  
 |||||:|||||  
 Db 415 ThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThrValcylHisPhe 434  
 QY 1171 GCGTTTACCAAGTACGCGCTACCTCGACTGATCCATGCGCAGATCAGAGACAAG 1230  
 |||||:|||||  
 Db 435 GlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeuMetArgSerGlu 454  
 QY 1231 GAAGCCCC 1239  
 |||||  
 Db 455 ProArgPro 457  
 RESULT 10  
 Q9GMD9 PRELIMINARY; PRT; 469 AA.  
 ID Q9GMD9  
 AC Q9GMD9  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Coagulation factor X.  
 OS Ornithorhynchus anatinus (Duckbill platypus).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.  
 OX NCBI\_TaxID=9258;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21015017; PubMed=11132153;  
 RA Poorafshar M., Aveskogh M., Munday B., Hellman L.;  
 BA "Identification and structural analysis of four serine proteases in a  
 L monotreme, the platypus, Ornithorhynchus anatinus.";  
 CC Immunogenetics 52:19-28(2000).  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AF275654; AAG00453.1; -.  
 DR HSPF; P00742; 1XKB.  
 DR MEROPS; S01.216; -.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF-2.  
 DR InterPro; IPR001881; EGF-Ca.  
 DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF\_like; 2.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; TRYP\_SFC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.

DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 469 AA; 52196 MW; 4C66C230D0758F6A CRC64;  
 Alignment Scores:  
 Pred. No.: 4,39e-48 Length: 469  
 Score: 773.00 Matches: 162  
 Percent Similarity: 53.27% Conservative: 74  
 Best Local Similarity: 36.57% Mismatches: 157  
 Query Match: 33.08% Indels: 50  
 DB: Gaps: 10  
 US-09-763-153-1 (1-1245) x Q9GMD9 (1-469)  
 QY 1 GCCAACTCTTCCTGGAGAGCTCCGTACACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
 |||||:|||||  
 Db 41 AlaAsnSerLeuPheGluLeuLysLysGlyAsnLeuGluArgGluCysAsnGluGlu 60  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACTGGCCTTC 120  
 |||||:|||||  
 Db 61 ThrCysSerTyrGluGluAlaArgGluValPheGluAspThrAspLysThrAsnGluPhe 80  
 QY 121 TGTCTCAAGCAGCTGCGAGCTGACCTGTGCTTGTGCTTGGCCTGGAGCACCCCTGGGCC 180  
 |||||:|||||  
 Db 81 TrpAsnIleTyrLysAspGlyAsnGlnCysGluThrGlnPro----- 94  
 QY 181 AGCCTGTGCTGGCGGCGACGCTGCATGACAGCATCGCAGCTTCAGCTCGCACTGC 240  
 |||||:|||||  
 Db 95 -----CysGlnAsnGlnGlyValCysLysAspGlyLeuAlaGluTyrThrCysLysCys 112  
 QY 241 CGCAGCGCTGGGAGCGCTCTTCCAGCGGAGGTGAGCTTCCTCAATTGCTCGCTG 300  
 |||||:|||||  
 Db 113 SerAlaGlyTyrGluGlyLysAsnCysAspThrThrValLysIle---CysSerLeu 131  
 QY 301 GACAAACGCGCTGTCACGCATTACTGCCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360  
 |||||:|||||  
 Db 132 AsnAsnGlyAspCysGluGlnPheCysLysSerValValAsnThrValValCysSerCys 151  
 QY 361 GCGCTCGCTACAGCTGGGAGGAGCTCCTGCAGCTGCACCCGCGAGTGAAGTTCCT 420  
 |||||:|||||  
 Db 152 AlaGlnGlyTyrIleLeuGlyAspGlnLysSerCysIleProThrValProPhePro 171  
 QY 421 TGTGGAGGCGCTCGAAGCGATGGAGAAGAAG----- 453  
 |||||:|||||  
 Db 172 CysGlyLysLeuThrValGlyArgLysArgSerArgGluLeuProGluGluGlnAsp 191  
 QY 454 -----CCAGTACCTGAAACGAGAC-----ACAGAAGACCAAGAACCA 495  
 |||||:|||||  
 Db 192 GlyAspAsnAlaHisValAlaGluAspValLeuGluAlaThrGluAsnProPheGluAla 211  
 QY 496 GTAGATCCG-----CGGCTC 510  
 |||||:|||||  
 Db 212 GluProProAspGlnAsnThrThrLeuAlaGluProGlyGluAsnAlaLeuValArgIle 231  
 QY 511 ATTGATGGGAAGATGACACGCGGGGAGACAGCCCTCGGAGGTGGTCCCTGCTG---GAC 567  
 |||||:|||||  
 Db 232 ValcylGlyArgGluCysHisAspGlyGlyCysProTrpGlnAlaLeuValAsnAsp 251  
 QY 568 TCRAAGAAGAAGCTGGCTCGCTCATCCACCCCTCTGGTGGTGGTGGAGCG 627  
 |||||:|||||  
 Db 252 GluAsnGlyGlnGlyPheCysGlyGlyThrIleLeuAsnGluTyrIleLeuSerAla 271  
 QY 628 GCCCACTGCATGATGATGATCCAGAGACTCCTTGTGAGCTGGAGATGACCTCGCG 697  
 |||||:|||||  
 Db 272 AlaHisCysMethHisGlnAlaLysValArgPheLysValArgValGlyGluArgAspThrGlu 291  
 QY 688 CGCTGGGAGAAGTGGGAGCTGGACCTGGACATCAAGAGGTCTTCGTCCACCCCACTAC 747

```
Db 292 LysLysAspSerSerGluMetAlaHisGluValGluLysValIleValHisSerLysPhe 311
    ::: ::: |||::: ::::::::::::::||| ||||| :::
Qy 748 AGCAGAGCACCACCGCATGACATCGCTGCTGACCTGGCCAGCGCCGACCCCTC 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 ValLysLysThrThrAspPheAspIleAlaValIleLysLysThrProIleThrPhe 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 808 TCGCAGACCATAGTCGCCATCTGCTCCGAGCAGCGCGCTTCGAGAGCGCGAGCTC--- 864
    ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 ArgMetAsnValSerProAlaCysLeuProGluLysAspTrpAlaGluAspIleLeuMet 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 865 ---AATCAGCGCGCGCAGGACCCCTGCTGACGGGCTGGGCTACCCAGCAGCGCGAGAG 921
    ::||| ||| ::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 352 AsnGlnLysAlaGly-----ValValSerGlyPheGly-----ArgVal 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 922 AAGGAGCGCAAGAAACCGCACCTGCTCTCACTTCATCAAGATTCGCGTGGTCCCG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 HisGluLysGlyArgProSerThr---ValLeuLysMetLeuGluValProTyrValGlu 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 982 CACATGAGTCAGCGAGGTGTCATGAGCAACATGCTGTGAGAACATGCTGTGCGGGC 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 ArgThrThrCysLysGlnSerSerSerPheAspIleThrProAsnMetPheCysAlaGly 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1042 ATCTCGGGGACCGCAGGATGCTGGAGGCGCACAGTGGGGGCCCATGCTGCTCC 1101
    ::||| ||| ::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 404 TyrAspSerArgProGluAspAlaCysGlnGlyAspSerGlyGlyProHisValThrLys 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1102 TTCCAGCGCACCTGCTCTGCTGGCTGCTGAGCTGGGTGAGGCTGTGGGCTCCTT 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 TyrLysAspThrThrPheValThrGlyIleValSerTrpGlyGluGlyCysAlaGlnAsn 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1162 CACAACACTACGGCGTTACACAAAGTCAGCGCTACCTCGACTGATCCATCGGCACATC 1221
    ::||| ||| ::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 444 GlyLysPheGlyValTyrThrLysAlaAlaThrPheLeuSerTrpIleLysArgMetMet 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1222 AGAGCAAG 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 ArgGlnLys 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
Q61109 PRELIMINARY; PRT; 446 AA.
ID Q61109;
AC Q61109;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor VII.
GN F7 OR FVII.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
TISSUE=LIVER.
MEDLINE=96276538; PubMed=8701412;
RA Idusogie E., Rosen E., Geng J.P., Carmeliet P., Collen D.,
RA Castellino F.J.;
RT "Characterization of a cDNA encoding murine coagulation factor VII.";
RL Thromb. Haemost. 75:481-487(1996).
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U4795; AAC52570.1; -.
DR HSSP; P08709; IFAK.
DR MEROPS; S01.215; -.
DR MGD; MGI:109325; F7.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_blood.
```

```
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; Vitk_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_Like; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLL_CARBOXYLATION; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW Serine protease.
SQ SEQUENCE 446 AA; 50318 MW; 482FD09BEFDA6870 CRC64;
```

```
Alignment Scores:
Pred. No.: 5,15e-48 Length: 446
Score: 772.00 Matches: 163
Percent Similarity: 56.43% Conservative: 74
Best Local Similarity: 38.81% Mismatches: 147
Query Match: 33.03% Indels: 36
DB: 11 Gaps: 11
```

US-09-763-153-1 (1-1245) x Q61109 (1-446)

```
Qy 1 GCACATCTCTCTGAGGAGCTCCGTCACAGACGCTGAGCGGAGTCATAGAGAG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 AlaAsnSerLeuLeuGluLeuTrpProGlySerLeuGluArgGlyCysAsnGluGlu 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 GlnCysSerPheGluGluAlaArgGluIlePheLysSerProGluArgThrLysGlnPhe 81
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 TGTCTCAACACGCTCGACGCTGACCATGCTTGGTCTTCCCTTGGAGCACCGCTGGCC 180
    ||| ::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 TrpIleValTyrSerAspGlyAspGlnCys-----AlaSerAsnProCysGln 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 AGCCTGTGCTGGCGGACGCGACGCTCATCGAGCATCGGAGCTTCAGCTCGCAGCTC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 98 AsnVal-----GlyThrCysGlnAspHisLeuLysSerTyrValCysPheCys 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 CGCAGCGCTGGAGGGCGCTTCTCCAGCGGGAGGTGAGC---TTCCTCAATGCTCG 297
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 LeuLeuAspPheGluGlyArgAsnCysGluLysSerLysAsnGluLeuIleCysAla 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 298 CTGACAGCGCGGCTGCACGCTTACTGCTAGAGAGGTGGCTGGCGGCGC---TGT 354
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 AsnGluAsnGlyAspCysAspGlnTyrCysArgAspHisValGlyThrLysArgThrCys 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 355 AGCTGTGCGCCTGGCTCAAGCTGAGGAGGAGCTCTCTGCTGAGTGCACCCGACGTGAAG 414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 SerCysHisGluAspTyrThrLeuGlnProAspGluValSerCysLysProLysValGlu 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 415 TTCCTTGTGGGAGGCCCTGGAGAGCGGATGGAGAGACGCGAGCTCCTGAAACAGAC 474
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 TyrProCysGlyArg---IleProValValGluLysArgAsnSerSerArgGlnGly 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 475 ACAGAGACCAAGAACAGCAAGTAGATCCGCGCTCATTTGATGGGAGAGATGACCGAGCG 534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 -----ArgIleValGlyCysAsnValCysProLys 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 535 GGAGACAGCCCTGGCAGGTGGTCTGCTGAGCTCAAGAGAGAGCTGCGCTCGCGGGCA 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 GlyGluCysProTrpGlnAlaVal---LeuLysIleAsnGlyLeuLeuLysGlyAla 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
QY 610 TCCTGGTGTGACAGGGCCCACTGATGATGACTCAAGAAGCTCCTGTGAGGCTT 669
Db 610 LysTrpIleValThrAlaAlaHisCysValAspThrGlyValIleThrValAla 279
QY 670 GGAGATGATGACCTGGCGCTGGGAGAGTGGGAGTGGACCTGGACATCAAGAGGTC 729
Db 670 GlyGluHisAsnIleGluGluThrGluHisThrGluGluLysArgAsnValIleArgIle 299
QY 730 TTGCTCCACCCCACTACACAGACACACCACC-----GACATGACATCGCAGTGG 783
Db 730 IleProHisAsnIleGluGluThrGluHisThrGluGluLysArgAsnValIleArgIle 319
QY 784 CACCTGGCCAGCGCCGACCTCTCGCAGACCATAGTCCCTGCTCCCGGACAGC 843
Db 784 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 339
QY 844 GGCCTTCGAGCGCGGAGCTCAATCAGCGCGGCGGACGACCTCTGAGCGGCTGGGC 903
Db 844 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTyrGly 357
QY 904 -----TACCACAGCAGCCGACGAGAGGAGGCGCAAGAAACCGACCTTCGCTCAAC 957
Db 904 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370
QY 958 TTATCAAGATTCCTGGTGGTCCGACAAATGAGTGCAGCGAGTCAATGACCAACATGGTG 1017
Db 958 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 390
QY 1018 TCTGAGACATGCTGTGGGGGATCTCTCGGGGACCGGACGATGCTCGAGGCGGAC 1077
Db 1018 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410
QY 1078 AGTGGGGGGCGATGCTGCTCTCCACGCGACCTGCTGCTGGTGGCGCTGGTGGC 1137
Db 1137 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430
QY 1138 TGGGGTGGAGGCTGTGGGCTCTTCAACACTAGCGGCTTTACACCAAGTCAAGCGCTAC 1197
Db 1138 TrpGlyGluGluCysAlaMetLysGlyTyrGlyIleTyrThrLysValSerArgTyr 450
QY 1198 CTCGACTGGATC 1209
Db 1198 ValAsnTrpIle 454

RESULT 13
Q14316 PRELIMINARY; PRT; 456 AA.
AC Q14316;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE F9 (Coagulation factor IX (Plasma THROMBOPLASTIC component, christmas disease, HAEMOPHILIA B)) (Factor IX).
F9 OR FACTOR IX.
F9 Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 3-19 FROM N.A.
RX MEDLINE=88327116; PubMed=3416069;
RA Rietsma P.A., Bertina R.M., Ploos van Amstel J.K., Riemsma A.,
RA Briet E.;
RT "The putative factor IX gene promoter in hemophilia B Leyden.";
RL Blood 72:1074-1076(1988).
CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AL033403; CAA21954.1; -.
DR EMBL; X55008; CAB38245.2; -.

```

```
DR HSP; P00740; 1CFH.
DR MEROPS; S01.214; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFBLD.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW Serine protease.
SQ SEQUENCE 456 AA; 51149 MW; 54E20A1B3964E234 CRC64;

Alignment Scores:
Pred. No.: 2,16e-45 Length: 456
Score: 736.00 Matches: 150
Percent Similarity: 52.36% Conservative: 72
Best Local Similarity: 35.38% Mismatches: 156
Query Match: 31.49% Indels: 46
DB: 4 Gaps: 10

US-09-763-153-1 (1-1245) x Q14316 (1-456)
QY 13 CTGGAGGAGCTCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGAGATCTGTGACTTC 72
Db 47 LeuGluGluPheValGlnGlyAsnLeuGluArgGluCysMetGluGluLysCysSerPhe 66
QY 73 GAGGAGGCCAAGAAATTTTCAAAATGTGATGACACACTGGCCTCTGTGTCACAGCAC 132
Db 67 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGlnTyr 86
QY 133 GTGCGAGGTGACAGTGTGCTGTGCTGCTGGCCCTTGGAGCACCCCTGGCGGCTGCTGC 192
Db 87 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 98
QY 193 GGGCAGCGGCGCTGTCATCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
Db 99 AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 118
QY 253 GAGGCGCGCTTGTGCGCGCGCGGAGGTGAGCTTCTCCTCAATGTCTGCGTGGACACGCGGC 312
Db 119 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 135
QY 313 TGCACCATCTACTGCTCTAGAGGAGGTGGCTGGCGG---CGCTGTAGCTGTGGCTGCTGC 369
Db 136 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 155
QY 370 TACAAGCTGGGGAGCAGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
Db 156 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 175
QY 430 CCTGGAAGCGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 471
Db 176 ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProAspValAsp 195

```

```
QY 472 -----GACAGAAGAACCAAGAACCAAGTAGAT----- 501
Db 196 TyrValAsnSerThrGluAlaGluThrIleLeuAsnIleThrGlnSerThrGlnSer 215
QY 502 -----CCGGGCTCATTGATGGGAAGATGACCAAGGCGGGGACAGACCCCTGG 549
Db 216 PheAsnAspPheThrArgValValGlyGlyGluAspAlaLysProGlyGlnPheProTyr 235
QY 550 CAGGTGCTCTGTCGACCAAGAACAGAGCTGGCTGGGGCAGTGTCTATCCACCC 609
Db 236 GlnValVal-----LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu 254
QY 610 TCCTGGTGTGACAGCGGCCACTGATGATGAGTCAAGAGAGCTCTTGTTCAGGCTT 669
Db 255 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValValAla 274
QY 670 GGAGATGATGACCTGGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGAGGTC 729
Db 275 GlyGluHisAsnIleGluGluThrGluHisThrGluGluLysArgAsnValIleArgIle 294
QY 730 TTCGTCCACCCCACTACAGCAAGACACACC-----GACAATGACATCCGACTGTG 783
Db 295 IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu 314
QY 784 CACCTGGCCCGCCCGCCCTCTCGCACCATAGTCCCATCTGCTCCCGGACAGC 843
Db 315 GluLeuAspGluProLeuValLeuAsnSerThrValThrProIleCysIleAlaAspLys 334
QY 844 GGCTTGGCAGCGCGAGCTCAATCAGCGCGCCAGGAGACCTCGTGACGGGCTGGGC 903
Db 335 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 352
QY 904 -----TACCACAGCAGCGAGAGAGGAGGCCACAGAAACCGACCTTCCTCTCAAC 957
Db 353 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 365
QY 958 TTCATCAAGATTCCTCGTGTCCCGCAATGAGTGCAGCGGAGTTCATGAGCAACATGGT 1017
Db 366 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 385
QY 1018 TCTGAGAATGCTGTGTGGGGCATCTCCGGGAGCCGCGAGATGCTTCGAGGCGAC 1077
Db 386 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyGlyArgAspSerCysGlnGlyAsp 405
QY 1078 AGTGGGGGCCATGCTGCTCCCTCTCCACGCGACCTGGTTCCTGGTGGGCTGTGTGAC 1137
Db 406 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleSer 425
QY 1138 TGGGTGAGGGGCTGTGGCTCTTCAACACTACGGCGTTTACACCAAGTCAAGCGCTAC 1197
Db 426 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 445
QY 1198 CTCGACTGCATC 1209
Db 446 ValAsnTrpIle 449
RESULT 14
Q95ND6 PRELIMINARY; PRT; 461 AA.
AC Q95ND6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Coagulation factor XI.
GN F9.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=505;
```

```
RA Satta Y.;
RT "Comparison of DNA and protein polymorphisms between humans and
RL chimpanzees.";
DR EMBL; AB062471; BAB58886.1; -.
DR EMBL; AB062459; BAB58886.1; JOINED.
DR EMBL; AB062461; BAB58886.1; JOINED.
DR EMBL; AB062463; BAB58886.1; JOINED.
DR EMBL; AB062465; BAB58886.1; JOINED.
DR EMBL; AB062467; BAB58886.1; JOINED.
DR EMBL; AB062469; BAB58886.1; JOINED.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 461 AA; 51695 MW; 8F5A69A525DF65B5 CRC64;

Alignment Scores:
Pred. NO.: 4.41e-44 Length: 461
Score: 718.00 Matches: 149
Percent Similarity: 50.94% Conservative: 67
Best Local Similarity: 35.14% Mismatches: 162
Query Match: 30.72% Indels: 46
DB: Gaps: 10

US-09-763-153-1 (1-1245) x Q95ND6 (1-461)
QY 13 CTGAGGAGCTCCGTACAGCAGCGCTGGAGGGAGTGTCATAGAGAGATCTGTGACTTC 72
Db 52 LeuGluGluPheValGlnGlyAsnLeuGluArgGluCysMetGluGluLysCysSerPhe 71
QY 73 GAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCTTCTGTGTCCAAAGCAC 132
Db 72 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGluTyr 91
QY 133 GTCGACGGTGACCACTGCTGTGCTTGTCCCTTGGACGACCGCTGGCGGAGCTGTGTCG 192
Db 92 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 103
QY 193 GGGCAGCGCAGCTGATCGACGCGCTTACGCTGCGACTGCGGAGCGGCTGG 252
Db 104 AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 123
QY 253 GAGGGCGCTTCTCCAGCGCGGAGGTGCTTCTCAATTCGCTGCTGGACAAAGCGGCG 312
Db 124 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 140
QY 313 TGCACGCATTACTGCTTAGAGGAGGTGGCTGGCGG---CGCTAGCTGCTGGCCCTGGC 369
Db 141 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 160
QY 370 TACAAGCTGGGGGACGACCTCTCTGCTGCTACCCCGCAGTCAAGTTCCTTGTGGGAGG 429
Db 161 TyrArgLeuAlaGluAsnGlnLysSerGluProAlaValProPheProCysGlyArg 180
QY 430 CCCTGGAGCGGATGGAGAGAGCGGACGTCACCTGAAA----- 468
Db 181 ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProAspValAsp 200
```

```

QY 469 -----CGAGACACAGAA-----GACCAAGAA 489
Db 201 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnThrProGlnSer 220
QY 490 GACCAAGTAGATCCCGGCTCATTCATGGAGATGACCGGGGGAGACAGCCCTGG 549
Db 221 *****ThrArg***ValGlyGlyGluAsp***AsnProGlyGlnTyrPro--- 239
QY 550 CAGGTGGTCTCTGCAGCTCAAGAGAGCTGGCTCGGCGAGTCTCTCATCACCCC 609
Db 240 Arg***ValValLeuAsnGlyLysValAspAlaPheCysGlySerIleValAsnGlu 259
QY 610 TCTGTGGTCTGACAGCGCCCACTGCATGATGAGTCCAAAGAGCTCTTGTGAGGCT 669
Db 260 LysTrpIleValAlaAlaHisCysValAspThrGlyValLysIleThrValValAla 279
QY 670 GGAGAGTAGACTCGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGAGGTC 729
Db 280 GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 299
QY 730 TTCGTCCACCCCACTACAGCAAGAGACACC-----GACAATGACATCGCATGCTG 783
Db 300 IleProHisHisAsnTyrAsnAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu 319
QY 784 CACGTGGCCACCCGCCACCTCTCGCAGACCATAGTCCCATCTGCTCCCGGACAGC 843
Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 339
QY 844 GGCCTTGCAGACGGGAGCTCAATCAGCGCGCCAGGAGACCTCTGACGGCTGGGGC 903
Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerglyTyrValSerGlyTrpGly 357
QY 904 -----TACCACGACGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
Db 358 ArgValPheHisGlyArgSer-----AlaLeuValLeuGln 370
QY 958 TTCATCAAGATTCCTGGTCCCGCACAATGATGTCACGAGGTCATGAGCAACATGGTG 1017
Db 371 TyrLeuArgValProLeuValAspAlaThrCysLeuArgSerThrLysPheThrIle 390
QY 1018 TCTGAGAACATCTGTGGCGGCTCTCGGGACCGGAGGAGGAGGAGGAGGAGGAGGAG 1077
Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410
QY 1078 ATGGGGGGCCCATGCTGCCCTCCCTTCCACGCGACCTGCTGCTGGTGGCTGGTGAGC 1137
Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430
QY 1138 TGGGGTGAGGGCTGGGGCTCTTCCACTACGCGGTTTACACCAAGTACGCGGCTAC 1197
Db 431 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 450
QY 1198 CTCGACTGGATC 1209
Db 451 ValAsnTrpIle 454

RESULT 15
Q9PTW7 PRELIMINARY; PRT; 608 AA.
AC Q9PTW7
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Prothrombin.
GN OSPT.
OS Struthio camelus (Ostrich).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC Struthio.
OX NCBI_TaxID=8801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;

```

```

RX MEDLINE=20579470; PubMed=11137455;
RA Frost C., Naude R., Oelofsen W., Muramoto K., Naganuma T., Ogawa T.;
RT "Purification and characterization of ostrich prothrombin.";
RL Int. J. Biochem. Cell Biol. 32:1151-1159(2000).
CC -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB028871; BAA89046.1; -.
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR023833; GLA_blood.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9AE54EA2 CRC64;

Alignment Scores:
Pred. No.: 5,15e-30 Length: 608
Score: 525.00 Matches: 139
Percent Similarity: 39.00% Conservative: 59
Best Local Similarity: 28.44% Mismatches: 175
Query Match: 22.46% Indels: 166
DB: 13 Gaps: 19

US-09-763-153-1 (1-1245) x Q9PTW7 (1-608)
QY 10 TTCCTGAGGAGTCTCTCAGCAGCCTCGAGCGGAGTGCATAGAGGAGATCTGTGAC 69
Db 48 PheLeuGluMetLeuLysGlyAsnLeuGluArgGluCysLeuGluIleCysIle 67
QY 70 TTCGAGAGGCCAAGGAATTTCCAAATGTGGATCACACACTGGCCTTCTGTCTCAAG 129
Db 68 TyrGluGluAlaPheGluAlaLeuGluSerThrAlaArgThrGluGluPheTrpSerLys 87
QY 130 CAC----- 132
Db 88 TyrGlnAlaCysGlnGlyIleArgLysSerArgThrValLeuAspAlaCysLeuGluGly 107
QY 133 -----GTGACGGT 141
Db 108 AsnCysSerThrAspLeuGlyGlnAsnTyrArgGlyThrIleSerHisThrLysSerGly 127
QY 142 GACCACTGCTGGTCTTCCCTTGGACACCCG----- 174
Db 128 ThrGluCysGlnMetTrpThrSerLysTyrProHisIleProLysPheAsnThrIle 147
QY 174 ----- 174
Db 148 HisProAsnLeuIleGluAsnTyrCysArgAsnProAspAsnAsnProGluGlyProTrp 167
QY 175 -----TGCAGCCAGCTGTCTCGGGGCAC 198
Db 168 CysTyrThrArgAspProThrValProArgGluCysProIleProValCysGlyGlu 187
QY 198 ----- 198

```



Db 188 GluArgThrThrValGluPheThrProArgValLysProProAlaSerThrGluProCys 207  
QY 199 -----GGCAGGTCATCGAGCGCATCGGACGCTTCAGC 231  
Db 208 GluGlnGluLysGlyMetLeuTyrAlaGlyThrLeuSerValThrIleSerGlyAlaLys 227  
QY 232 TGC-----GACTGCCGCGAGCGGCTGGAGGGCGCTTCGCCAGCGC 273  
Db 228 CysLeuProTrpAsnSerGluLysAlaLysGluValLeuGlnGlyLysThrIleLeuThr 247  
QY 274 GAGGTGAGCTTCCTCAAT----- 291  
Db 248 GluValLysLeuLeuGluAsnTyrCysArgAsnProAspAlaAspAspGluGlyValTrp 267  
QY 292 TCGTCGCTGGACAAGCGGGCTGCACGATTAATCTGC----- 327  
Db 268 CysValThrAspGluProHisPheGluTyrCysAspLeuGlnTyrCysAspSerSer 287  
QY 328 -----CTAGAGGAGGTGGCTGGCGCGCTGTAGTGTGGCGCT 366  
Db 288 LeuGluAspGlyAsnGluGlnLeuGluGluIleSerGlyArg-----ThrValLeuGln 305  
QY 367 GGCCTCAAG---CTGGGGAGCAGCTCTCGAGTGTCAACCGCGCAGTGAAGTTCCTTGT 423  
Db 306 GluTyrLysThrPhePheAsp-----LysThrPheGlySerGlyGluAlaAspCys 323  
QY 424 GGG---AGCCCTGAGAGGGATGAGAGAGAGCCAGCTCACTCAAGAGAGACAGAA 480  
Db 324 GlyIleArgPro-----LeuPheGluLysLysLysIleLysAspLysSerGluLysGlu 341  
QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATGTGGAAGATGACAGCGGGGAGAC 540  
Db 342 LeuLeuGluSerTyrIleGlySerArgValValHisGlyAspAlaGluValGlySer 361  
QY 541 AGCCCTGCGCAGGTGCTCTGCTG---GACTCAAGAGAGAGAGCTGGCTGGCGGCGAGT 597  
Db 362 SerProTrpGlnValMetLeuTyrLysLysSerProGlnGluLeuLeuCysGlyAlaSer 381  
QY 598 CTCATCCACCCCTCTGCTGGTGTGACAGCGGCCACTGCATG-----CAT 642  
Db 382 LeuIleSerAspSerTrpValLeuThrAlaAlaHisCysLeuTyrTyrProTrpAsp 401  
QY 643 GAG-----TCCAAGAAGCTCTCTGAGGCTTGAGAGATGATGACCTGCGGCGTGG 693  
Db 402 LysAsnLeuThrThrAsnAspIleLeuValArgIleGlyLysHisPheArgAlaLysTyr 421  
QY 694 GAGAGTGGGAGCTGAGCTGAC---ATCAAGAGAGTCTGCTGCCACCCCACTACAGC 750  
Db 422 GluLysAsnLysGluLysIleAlaLeuLeuAspLysIleIleHisProLysTyrAsn 441  
QY 751 ---AAGAGCACCACGACAATGACATCGACTGCTGCGACCTGGCCCGCCGCCACCCCTC 807  
Db 442 TrpLysGluAsnMetAspArgaspIleAlaLeuMetHisLeuLysArgProValIlePhe 461  
QY 808 TCGCAGACCATAGTCCCATCTGCTCCCGGACAGCGGCTTGCAGAGCGGAGTCAAT 867  
Db 462 SerAspTyrIleHisProValCysLeuProThrLysGluValAlaGlnArg---LeuMet 480  
QY 868 CAGCGCGCGCAGGAGACCTCTGACGGCTGGGGCTACACAGCAGCGCGAGAGAGAG 927  
Db 481 LeuAlaGlyTyrLysGlyArgValThrGlyTrpGluLeuLysGluThrTrpAlaThr 500  
QY 928 GCCAAGAGAACCGCACCTTCGCTCACTCACTCATCAAGATCCCGTGGTCCCGCACAT 987  
Db 501 AsnProSerAsnLeuProThrValLeuGlnGlnLeuAsnValProIleValAspGlnAsp 520  
QY 988 GAGTGCAGCGAGGTGATGACGAACATGCTGTGAAACATGCTGTGCGGCGATCCTC 1047  
Db 521 ThrCysLysAlaSerThrLysValLysValThrAspAsnMetPheCysAlaGlyTyrSer 540  
QY 1048 GGGGAC-----CGGCAGGATGCTGCGAGGCGCAGTGGGGGCCCATGCTC--- 1095  
Db 541 ProGluAspSerLysArgGlyAspAlaCysGluLysAspSerGlyGlyProPheValMet 560

QY 1096 ---GCCTCCTCCACGCGACCTGCTTCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGT 1152  
Db 561 LysSerProAspAspSerArgTyrTyrGlnValGlyIleValSerTrpGlyGlyCys 580  
QY 1153 GGGCTCCTCCACAACTACGCGCTTTACACCAAAAGTCAGCGCTACCTCGACTGGATC 1209  
Db 581 AspArgAspGlyLysTyrGlyPheTyrThrHisValPheArgLeuLysLysTrpIle 599  
RESULT 16  
Q91001 PRELIMINARY; PRT; 607 AA.  
ID Q91001 AC Q91001;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Thrombin.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=92212913; PubMed=1557383;  
RA Banfield D.K., Macgillivray R.T.;  
RT "Partial characterization of vertebrate prothrombin cDNAs:  
RT amplification and sequence analysis of the B chain of thrombin from  
RT nine different species.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=94233694; PubMed=7513365;  
RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;  
RT "Evolution of prothrombin: isolation and characterization of the cDNAs  
RT encoding chicken and hagfish prothrombin.";  
RL J. Mol. Evol. 38:177-187(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX Banfield D.K.;  
RA Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
EMBL; M81391; AAA21619.1; -;  
DR HSP; P00734; IUVS.  
DR MEROPS; S01.217; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR002383; GLA\_blood.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003966; Prothrombin.  
DR InterPro; IPR001254; Ser\_protease\_try.  
DR InterPro; IPR000294; VitK\_dep\_GLA.  
DR Pfam; PF00594; gla; 1.  
DR Pfam; PF00051; kringle; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00001; GLABLOOD.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 2.  
DR SMART; SM00069; GLA; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS00070; KRINGLE\_2; 2.  
DR PROSITE; PS00240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
KW Hydrolyase; Serine protease.  
SQ SEQUENCE 607 AA; 69110 MW; 002F3605EA36270F CRC64;







Db 622 ArgLeuAsnLysProGlyValThrArgValAlaAsnTyrValAspTyrIleAsnAsp 641  
 QY 1216 CACATCAGACAAAGGAGCCGCC 1239  
 Db 642 ArgIleArgProLysArgPro 649  
 RESULT 19  
 Q9CQW3  
 AC Q9CQW3; PRELIMINARY; PRT; 399 AA.  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE 1300015B06Rik protein.  
 1300015B06Rik.  
 Mus musculus (Mouse).  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, AND LIVER;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC ENBL; AK008819; BAB25912.1; -.  
 CC ENBL; AK005011; BAB23744.1; -.  
 CC HSP; P00740; ICFH.  
 CC MEROPS; S01.996; -.  
 CC MGD; MGI:1914151; 1300015B06Rik.  
 CC InterPro; IPR000152; Asx\_hydroxyl.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR000561; EGF-like.  
 CC InterPro; IPR000742; EGF\_2.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR002383; GLA\_blood.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC InterPro; IPR000294; vitk\_dep\_GLA.  
 CC Pfam; PF00008; EGF; 2.  
 CC Pfam; PF00594; gla; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00001; GLABLOOD.  
 CC SMART; SM00179; EGF\_CA; 2.  
 CC SMART; SM00069; GLA; 1.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
 CC PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 CC PROSITE; PS00186; EGF\_2; 2.  
 CC PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 CC Hydrolase; Serine protease.

SQ SEQUENCE 399 AA; 44304 MW; 4FC85C9598F27E03 CRC64;  
 Alignment Scores:  
 Pred. No.: 4,85e-26 Length: 399  
 Score: 470.00 Matches: 126  
 Percent Similarity: 43.71% Conservative: 58  
 Best Local Similarity: 29.93% Mismatches: 143  
 Query Match: 20.11% Indels: 94  
 DB: 11 Gaps: 16  
 US-09-763-153-1 (1-1245) x Q9CQW3 (1-399)  
 QY 10 TTCCTGGAGGAGCTCCGTACAGCAGCGCTGGAGCGGAGTGCATPAGAGAGATCTGTGAC 69  
 Db 45 PheLeuGluGluIlePheGlnGlyAsnLeuGluLysGluCysTyrGluGluValCysAsn 64  
 QY 70 TTCGAGGAGGCAAGGAATTTCCAAATGTGTGATGACACACTGGCTTCTGTGTCAG 129  
 Db 65 TyrGluGluAlaArgGluValPheGluAsnAspValIleThrAspGluPheTyrArgGln 84  
 QY 130 CAGCTCGACGCTGACAGTGTCTGTGCTTGTGCTTGGAGCACCCTGCGCAGCCTGTGC 189  
 Db 85 TyrGlyGlyGlySerProCysValSerGlnPro-----Cys 96  
 QY 190 TCGGGGACGCGCAGCTGCATCGCAGCGCTGGAGTTCAGCTGCGAGTGGCGCAGCGC 249  
 Db 97 LeuAsnGlnGlyThrCysGluAspHisIleArgSerTyrSerCysThrCysSerProGly 116  
 QY 250 TGGAGGCGCGCTTCTGCCAGCGCGAGTGTGCTTCTCAATGCTCGCTGGAC---AAC 306  
 Db 117 TyrGluGlyLysThrCysAlaMetAlaLysAsn-----GluCysHisLeuGluArgThr 134  
 QY 307 GCGCGCTGCAGCAGTACTGCTAGAGAGTGGTGGCGGCTGTAGCTGTGGCGCT 366  
 Db 135 AspGlyCysGlnHisPheCysHisProGlyGlnSerSerTyrMetCysSerCysAlaLys 154  
 QY 367 GGCTFACAGTGGGGGAGCAGCTCTTCAGTGTACCCCGCAGTGAAGTTCCTGTGGG 426  
 Db 155 GlyTyrLysLeuGlyLysAspGlnLysSerCysGlyProSerAspLysCysAlaCysGly 174  
 QY 427 AGCCCTTGAAGCGGATGGAGAGAGCGAGTCACTGAAACGAGACACAGAACCA 486  
 Db 175 -----AlaLeuThrSerGluHisIle----- 181  
 QY 487 GAAGACCAAGTAGATCGCGGCTCATTTGATGGAAGATGACGAGCGGGGAGACAGC--- 543  
 Db 182 -----ArgMetThrLysSerSerGlnSerGln 190  
 QY 544 -----CCCTGGCAGGTGCTCTGCTGGAGTCAAAAGAAAGAGTGCCTGCGGGGCA 594  
 Db 191 ProSerPheProTyrGlnValArgLeuThrAsnSerGluGlyGluAspPheCysAlaGly 210  
 QY 595 GTCTCATCCACCCCTCTCGGTGTGACAGCGCGCCACTGCATGGATGAGTCCCAAGAAG 654  
 Db 211 ValLeuGluGlnGluAspPheValLeuThrThrAlaLysCysSerLeuLeuHisSerAsn 230  
 QY 655 CTCTTGTTCAGGCTTGGAGAGTATGACCTGCGCGCTGGAGAGTGGAGAGTGGAGCTG 714  
 Db 231 IleSerValLys-----AlaAsnValAspGlnArg-----Ile 241  
 QY 715 GACATCAAGGAGGTCTTCTGTCACCCCACTACACAAAGAGCACCACGACAATGACATC 774  
 Db 242 ArgIleLysSerThrHisValHisMetArgTyrAspGluGluSerGlyGluAsnAspVal 261  
 QY 775 GCACTGCTGACCTGGCCAGCGCCGCCACCTCTCTGCGAGACCATAGTGCCTGCTGCTC 834  
 Db 262 SerLeuLeuGlnLeuGluGluProLeuGlnCysProSerSerGlyLeuProValCysVal 281  
 QY 835 CCGGACAGCGGCTTTCAGAGCGCGAGCTCAATCAGCGCGCCGAGGAGACCTGCTGACG 894  
 Db 282 ProGluArgAspPheAlaGluHisValLeu---IleProGlyThrGluGlyLeuSer 300  
 QY 895 GGCTGGGCTACACAGCAGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954

```

Db 301 GlyTrp-----MetLeu 304
Qy 955 AACTTCATCAAGATCCCGTGGTCCG-----CACAA 987
Db 305 AsnGlyThrHisLeuAlaThrProMetLeuLeuSerValThrGlnAlaAspGlyGlu 324
Qy 988 GAGTGCAGCAGGATGATGAGCAACATGGTCTGTGAGAACATCTGTGT-----GCG 1038
Db 325 GluCysGlyGlnThrLeuAsnValThrValThrArgThrSerCysGluGlySer 344
Qy 1039 GGCATCTCCGGACCGGAGGAGTGTCTGGAGGGCGACAGTGGGGCCCATGTGTGCGC 1098
Db 345 ValValMetGlyProTrpValGlu-----GlySerValValThr 357
Qy 1099 TCCCTTCACAC---GGCACCTGTCTCTGGTGGGCTGGTGGAGCTGG-----GGT 1143
Db 358 ArgGluHisLysGlyThrTrpPheLeuThrGlyLeuLeuGlySerProProProGly 377
Qy 1144 GAGGGCTGTGGGCTCTCTCACACTAGCGGCGTTTACACCAAGTACGCGCTACCTCGAC 1203
Db 378 GlnSerGlnMetLeu-----LeuThrAlaValProArgTyrSerMet 392
Qy 1204 TGG 1206
Db 393 Trp 393

RESULT 20
Q9V4N6 PRELIMINARY; PRT; 1379 AA.
AC Q9V4N6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CG2105 protein.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Calniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Kaulush K., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kadir C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C., Li J., Li Z., Liang Y., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA SVIRSKAS R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "the genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL: AE003840; AAF59230.1; -.
CC HSP: P00763; IDPO.
CC FlyBase: FBgn0033192; CG2105.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000024; Fz_domain.
CC InterPro: IPR002172; LDL_recept_A.
CC InterPro: IPR001254; Ser_protease_Try.
CC InterPro: IPR001190; Srcr_receptor.
CC InterPro: IPR000217; Tubulin.
CC Pfam: PF01392; Fz; 1.
CC Pfam: PF00037; ldl_recept_a; 2.
CC Pfam: PF00089; tryptsin; 1.
CC PRINTS: PR00722; CHYMOTRYPsin.
CC PRINTS: PR00261; LDLRECEPTOR.
CC SMART: SM00063; FRI; 1.
CC SMART: SM00192; LDLa; 2.
CC SMART: SM00202; SR; 1.
CC SMART: SM00020; TRYp_SPC; 1.
CC PROSITE: PS00038; Fz; 1.
CC PROSITE: PS01209; LDLRA_1; 1.
CC PROSITE: PS00068; LDLRA_2; 2.
CC PROSITE: PS00287; SRCR_2; 1.
CC PROSITE: PS00240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC PROSITE: PS00227; TUBULIN; 1.
KW Glycoprotein; Hydrolase; Serine protease.
SQ SEQUENCE 1379 AA; 149489 MW; A593A9CC2167E4B7 CRC64;

Alignment Scores:
Pred. No.: 8,34e-24 Length: 1379
Score: 440.50 Matches: 133
Percent Similarity: 39.76% Conservative: 63
Best Local Similarity: 26.98% Mismatches: 174
Query Match: 18.85% Indels: 123
DB: 5 Gaps: 20

US-09-763-153-1 (1-1245) x Q9V4N6 (1-1379)
Qy 55 GAGGAGATCTGTGACTTCGAGGAGCCCAAGCAATTTTCCAAATGTGGATGACACACTG 114
Db 872 GluAspCysValGlyLeuAspGluValArgGluValMetArg----- 885
Qy 115 GCTTCTGTGTCACAGCAGCTGCGGTCGACCGTGCACGATGC-----TTGGTTCCTCCCTTG 165
Db 886 AlaAlaThrHisProLysCysAspGlyPheGlnCysAspGlnAsnArgCysLeuProGln 905
Qy 166 GAGCACCCCTGTC----- 177
Db 906 GluTyrValCysAspGlyHisLeuAspCysMetAspGlnAlaAspGluAlaLysCysGlu 925
Qy 178 -----GCCAGCTGTGTCGCGGACGCGCAGCTGCATC----- 210
Db 926 ArgCysGlyProAspGluIleTyrCysGlyAspSerGlnCysIleGlyThrLysHisIle 945
Qy 211 ---GAGCGATCGGACGCTTACGTCGCGACTCCCGCAGCGCTGGGAGGGCGCTTCTGC 267
Db 946 CysAspGlyIle-----IleAspCysProTyrGlyGlnAspGluArgAsnGly 961
Qy 268 CAGCGGAGGTGAGCTTCCTCAATTGCTCGTGGACACGCGGCTGCACCGCATTTACTGC 327

```



Db 962 LeuArg---LeuSerGluArgAsnGlyAspValGlyThrGlyValLeuGluValTyrArg 980  
QY 328 CTAGAGAGGTGGCTGGCGCGCTAGCTGTGCGCTGCTACAAG----- 375  
Db 981 IleGlyGlnArgGlnTrpMetPro---AlaCysValLysAsnTrpAspArgAlaValSer 999  
QY 376 -----CTGGGGCAGCACCTCTGCAGTGTCAACCCGCGAGTGAAG 414  
Db 1000 ProSerAlaValCysSerIleLeuGlyTyrSerAlaValAsnAlaThrSerValLeuThr 1019  
QY 415 TTCCCTTGTGGAGGCC-----TGAAGCGG 441  
Db 1020 GlnLeuThrHisArgProGluLeuAlaThrValAsnValSerThrAspIleTrpLysMet 1039  
442 ATGGAGAAAGCAGCGACCTGACCTGAAACGAGAC-----ACAGAAGAC 483  
1040 TyrAlaLysArgLysSerThrLeuMetGlnGluPheAlaAsnCysLysLysThrGluAsp 1059  
QY 484 CAA-----GAA 489  
Db 1060 TyrProMetAlaAspLeuThrCysSerAsnTyrGluCysGlyArgValLysArgGlyArg 1079  
QY 490 GACCAAGTAGATCCGCGCTCATTTGAGGAAGATGACCGCGGGAGAGAGCCCTGG 549  
Db 1080 HisLysProSerArgArgIleIleGlyThrGlnAlaSerProGlyAsnTrpProPhe 1099  
QY 550 CAGGTGGTCTGCTGAGCTCAAGAAAGAGCTG---GCCTGCGGGGCGAGTGTCTCATCCAC 606  
Db 1100 LeuAlaIleLeuGlyGlyProGluLysIlePheTyrCysAlaGlyValLeuIleSer 1119  
QY 607 CCTCTCTGGTGTGACAGCGCCACTGCTGATGATGAGTCCAAAGAGCTCTGTGTCAGG 666  
Db 1120 AspGlnTrpValLeuThrAlaSerHisCys----- 1129  
QY 667 CTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGAGCTGGACCTG----- 714  
Db 1130 ValGlyAsnTyrSerValIleAspLeuGluAspTrpThrIleGlnLeuGlyValThrArg 1149  
QY 715 -----GACATCAAGAGAGTCTTCGTCCACCC 741  
Db 1150 ArgAsnSerPheThrTyrSerGlyGlnLysValLysValLysAlaValIleProHisPro 1169  
QY 742 AACTACAGACAGACACACCC---GACAAATCACATCCGCTGTGACCTGGCGCCAGCC 798  
Db 1170 GlnTyrAsnMetAlaIleAlaHisAspAsnAspIleAlaLeuPheGlnLeuAlaThrArg 1189  
QY 799 GCCACCTCTCGCAGACCATAGTCCCATCTGCTCCGACAGCGCCCTTGCAGAGCGC 858  
Db 1190 ValAlaPheHisGluHisLeuLeuProValCysLeuProProSerVal-----Arg 1207  
QY 859 GAGCTCAATCAGCGCGCGAGACCCCTGCTGACGGCTGGGCTACACAGCAGCGCA 918  
Db 1208 AsnLeuHis---ProGlyThrLeuCysThrValIleGlyTrpGly-----LysArgGlu 1224  
QY 919 GAGAGAGGCGCAAGACAAACCGACCTTCTGCTCATCAATTCATCAAGATTCCTGGTC 978  
Db 1225 AspLysAspProLysSerThrTyrGlnTyrIleValAsnGluValGlnValProIleIle 1244  
QY 979 CCGCACAAATGAGTCAGCAGGTCATGAGCAACATG---GTGCTCTGAGAACATGCTGTGT 1035  
Db 1245 ThrArgAsnGlnCysAspGluTrpLeuAspAsnLeuThrValSerGluGlyMetValCys 1264  
QY 1036 GCGGCAATCTCGGGACCGCGAGATGCTGCGAGGCGACAGTGGGGGCGCCATGGTC 1095  
Db 1265 AlaGlyPheAspAspGlyGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuLeu 1284  
QY 1096 GCCTCTTCCACGGC-----ACCTGTTCTGGTGGCTGGTGGTGGTGGTGG 1146  
Db 1285 CysProTyrProGlyGluLysAsnArgTrpPheValGlyIleValSerTrpGlyIle 1304  
QY 1147 GGCTGTGGTCTCTTCAACTACGGCGTTTACACCAAGTACCGCGCTACCTCGACTGG 1206  
Db 1305 MetCysAlaHisProArgLeuProGlyValTyrAlaAsnValGlnTyrValProTrp 1324

QY 1207 ATCCATGGCACAATFCAGAGACAAGAGCCCGCCAGAG 1245  
Db 1325 IleGlnGluGlnIleAlaLysHisSerArgProIleLys 1337

Search completed: March 17, 2003, 17:35:20  
Job time : 95.5 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:28:54 ; Search time 38 Seconds  
(without alignments)  
1927.976 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctcggagga.....acaaggagagccccccaggaag 1245

Listing table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_Spool/US09763153/runat\_11032003\_084248\_2035/app\_query.fasta\_1.1415  
-DB=Issued\_Patents\_AA -QPMF=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=60 -DOALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153 -RCGN\_1\_1\_21 @runat\_11032003\_084248\_2035 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	419	1	US-08-295-411-1
2	2298	98.3	419	2	US-08-955-471-1
3	2298	98.3	419	4	US-09-667-570A-3
4	2298	98.3	419	5	PCT-US92-10242-1
5	2298	98.3	460	2	US-08-756-506-2
6	2298	98.3	460	2	US-08-756-506-4
7	2298	98.3	461	6	5225537-2
8	2292	98.1	461	6	5460953-3
9	2266	97.0	461	6	5270178-17
10	2266	97.0	461	6	5270178-18
11	2260.5	96.7	460	6	5270178-13
12	2260.5	96.7	460	6	5270178-14

13	2255	96.5	410	4	US-09-065-872-1	Sequence 1, Appli
14	2255	96.5	410	4	US-09-667-570A-1	Sequence 1, Appli
15	2253.5	96.4	461	6	5270178-2	Patent No. 5270178
16	2247	96.1	409	4	US-09-065-872-2	Sequence 2, Appli
17	2247	96.1	409	4	US-09-667-570A-2	Sequence 2, Appli
18	2245.5	96.1	460	6	5270178-15	Patent No. 5270178
19	2218.5	94.9	262	1	US-07-720-189-1	Sequence 1, Appli
20	1393	59.6	262	1	US-07-720-189-1	Sequence 1, Appli
21	1367.5	58.5	261	6	5270178-19	Patent No. 5270178
22	1358.5	58.1	261	6	5270178-20	Patent No. 5270178
23	1328	56.8	250	4	US-08-944-483-51	Sequence 51, Appli
24	1320.5	56.5	261	6	5270178-21	Patent No. 5270178
25	1302.5	55.7	261	6	5270178-5	Patent No. 5270178
26	809	34.6	487	1	US-08-469-486-53	Sequence 53, Appli
27	809	34.6	487	2	US-08-469-486-53	Sequence 53, Appli
28	809	34.6	492	1	US-08-469-486-2	Sequence 2, Appli
29	809	34.6	492	2	US-08-469-486-2	Sequence 2, Appli
30	801.5	34.3	448	5	PCT-US92-10068-1	Sequence 1, Appli
31	799.5	34.2	448	1	US-08-295-411-3	Sequence 3, Appli
32	799.5	34.2	448	2	US-08-295-411-3	Sequence 3, Appli
33	799.5	34.2	448	5	PCT-US92-10242-3	Sequence 3, Appli
34	783	33.5	406	1	US-08-295-471-5	Sequence 5, Appli
35	783	33.5	406	2	US-08-295-471-5	Sequence 5, Appli
36	783	33.5	406	5	PCT-US92-10242-5	Sequence 5, Appli
37	783	33.5	444	1	US-08-475-845-2	Sequence 2, Appli
38	783	33.5	444	2	US-08-475-845-2	Sequence 2, Appli
39	783	33.5	444	2	US-08-660-289-2	Sequence 2, Appli
40	783	33.5	444	2	US-08-537-807-2	Sequence 2, Appli
41	783	33.5	444	2	US-08-537-807-2	Sequence 2, Appli
42	783	33.5	444	3	US-08-464-233-2	Sequence 2, Appli
43	783	33.5	444	4	US-09-189-607-2	Sequence 2, Appli
44	783	33.5	444	4	US-09-378-907-2	Sequence 2, Appli
45	783	33.5	444	5	PCT-US94-05779-2	Sequence 2, Appli
46	783	33.5	466	1	US-07-882-202A-4	Sequence 4, Appli
47	783	33.5	466	1	US-08-021-615A-4	Sequence 4, Appli
48	783	33.5	466	1	US-08-321-777-4	Sequence 4, Appli
49	783	33.5	466	4	US-09-009-217-14	Sequence 14, Appli
50	783	33.5	466	4	US-09-009-656-14	Sequence 14, Appli
51	783	33.5	466	5	PCT-US93-04493-4	Sequence 4, Appli
52	742	31.8	415	1	US-08-073-531B-1	Sequence 1, Appli
53	742	31.8	415	2	US-08-766-288-1	Sequence 2, Appli
54	742	31.8	437	1	US-08-487-037-2	Sequence 2, Appli
55	738.5	31.6	488	1	US-08-487-037-1	Sequence 1, Appli
56	736.5	31.5	461	6	5521070-2	Patent No. 5521070
57	736	31.5	461	3	US-08-742-877-2	Sequence 2, Appli
58	736	31.5	461	4	US-09-053-871A-21	Sequence 21, Appli
59	734	31.4	437	1	US-08-487-037-3	Sequence 3, Appli
60	731	31.3	415	1	US-08-295-411-2	Sequence 2, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-295-411-1  
; Sequence 1, Application US/08295411  
; Patent No. 5679639

; GENERAL INFORMATION:

; APPLICANT: Griffin, John H.

; APPLICANT: Westers, Rolf M.

; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and

; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods  
; TITLE OF INVENTION: for Inhibiting Coagulation

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Research Institute

; ADDRESS: Office of Patent Counsel, The Scripps

; STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/295,411  
FILING DATE: 22-AUG-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,989  
FILING DATE: 18-NOV-1991  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI263.0C1  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FEATURE:  
ANTI-SENSE: NO  
NAME/KEY: Region  
LOCATION: 1..157  
OTHER INFORMATION: /note= "Protein C Light Chain"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 158..169  
OTHER INFORMATION: /note= "Protein C Activation"  
OTHER INFORMATION: Peptide"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 170..419  
OTHER INFORMATION: /note= "Protein C Heavy Chain"  
US-08-295-411-1

Alignment Scores:  
Pred. No.: 3,34e-167 Length: 419  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x US-08-295-411-1 (1-419)

QY 1 GCCAACTCCTCTCGAGGAGCTCGTCCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
|||||  
1 AlaasnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCCCTC 120  
21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 40  
QY 121 TGGTCCAGCAGCTGCAGCGGTGACCAAGTGTCTTGGCTTGGCCCTGGAGCACCCCTGGCC 180  
|||||  
41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCGTGTGCTGGGGCAGCGACGTGCATCGACGCGATCGGCGAGCTTCAGCTGGCAGTGC 240  
|||||  
61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGGCGCGCTTCGCCAGCGCGAGGTGAGCTTCCTCAATGTCTCGGTG 300  
|||||  
81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACACGGCGGCTGCAGCGATTACTGCCTAGAGGAGTGGCTGGCGGCTGTAGCTGT 360  
|||||

Db 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCCTGGCTACAAGCTGGGGAGCAGACCTCTCGAGTGTCAACCCCGCAGTGAAGTCCCT 420  
|||||  
Db 121 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGGCCCTGGAGCGGATGGAGAAGACGCGAGTCACTGAAACGAGACACAGAA 480  
|||||  
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAAGACCAAGTATAGTCCGCGCTCATTTGATGGGAAGATGACACAGCGGGAGAC 540  
|||||  
Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTGGCAGGTGTCTCTGGTGGACTCAAAGAAGAGCTGGCCCTCGCGGCGAGTCTC 600  
|||||  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTGGTGTGTGACAGCGGCCACTGCATGGATGAGTCCAAGAGCTCTCT 660  
|||||  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspLysSerLysLeuLeu 220  
QY 661 GTCAGGCTGGAGAGTATGACCTGGCGCGCTGGGAGAAGTGGAGCTGGACCTGGACATC 720  
|||||  
Db 221 ValArgLeuGlyGluTyrAspLeuArgA-gTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGGAGGTCTTCCTCCACCCCACTACAGCAAGAGCACCACCAATGACATCGCACATG 780  
|||||  
Db 241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260  
QY 781 CTCACCTGGCCCGAGCCCGCCACCTCTCGCAGACCATAGTGCCTCATCTGCCCTCCGGAC 840  
|||||  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
QY 841 AGCGCCTTCCAGAGCGGAGCTCAATCAGCGCGCGCCAGGAGACCTCTCGTACGGGCTGG 900  
|||||  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGCAAGAAACCGCACCTTCGCTCAACTTC 960  
|||||  
Db 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGGTGGTCCCGCACAATGAGTGCAGCGAGGTGCATGACCAACATGTGTCT 1020  
|||||  
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGTGTGTGGCGCATCTCGGGGACCGCGGAGGATCGCTCGCAGGCGGACAGT 1080  
|||||  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCCCATGTGTGGCTCTTCACCGGACCTGTGTCTGTGGGCTGTGTGAGCTGG 1140  
|||||  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGCTCTTCACACTAGCGGCTTTACACCAAGTACAGCCCTACCTC 1200  
|||||  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
QY 1201 GACTGGATCCATGGGCACATCAGAGCAAGAAAGAGCCCGCCAGAG 1245  
|||||  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

# RESULT 2

US-08-955-471-1  
Sequence 1, Application US/08955471  
Patent No. 5968751  
GENERAL INFORMATION:  
APPLICANT: Griffia, John H.  
APPLICANT: Mesters, Rolf M.  
TITLE OF INVENTION: Serine protease-Derived Polypeptides and  
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods  
TITLE OF INVENTION: for Inhibiting Coagulation  
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Office of Patent Counsel, The Scripps  
 ADDRESSEE: Research Institute  
 STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92037

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/955,471  
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/295,411

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI263.0C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937

TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Region

LOCATION: 1..157

OTHER INFORMATION: /note= "Protein C Light Chain"

FEATURE:

NAME/KEY: Region

LOCATION: 158..169

OTHER INFORMATION: /note= "Protein C Activation

OTHER INFORMATION: Peptide"

FEATURE:

NAME/KEY: Region

LOCATION: 170..419

OTHER INFORMATION: /note= "Protein C Heavy Chain"

US-09-955-471-1

# Alignment Scores:

Pred. No.:	3,34e-167	Length:	419
Score:	2298.00	Matches:	415
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.33%	Indels:	0
DB:	2	Gaps:	0

US-09-763-153-1 (1-1245) x US-08-955-471-1 (1-419)

QY	1	GCACACTCTCTCGGAGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGCAG	60
Db	1	AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	20
QY	61	ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCCCTC	120
Db	21	IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe	40
QY	121	TGTCACACAGCTGCAGCGGTGCACAGTGCCTGTGCTTCCTTCGAGCAGCCCGTGC	180
Db	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCys	60

QY	181	AGCCTGTCTCGCGGACGCGCACGTGCTATCGACGCGCATCGCAGCTTCAGTGCAGTGC	240
Db	61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
QY	241	CGCAGCGCTGGAGAGCGCGCTTCTGCCAGCGCGAGGTGAGCTTCCTCAATGCTCGCTG	300
Db	81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
QY	301	GACAACGCGCGCTGCACGCATTACTGCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT	360
Db	101	AspAsnGlyGlyCysThrHisIleCysLeuGluGluValGlyTrpArgCysSerCys	120
QY	361	CGCGCTGGCTACAAGCTGGGGACGACCTCTCGCAGTGTTCACCCCGCAGTCAAGTTCCT	420
Db	121	AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro	140
QY	421	TGTGGAGCGCTGGAAGCGGATGGAGAAGCGCAGTCCACTGTAACGAGACACAGAA	480
Db	141	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	160
QY	481	GACCAAGAAGACCAAGTAGATCCGCGCTCATTCATGGAGAGATGACAGCGGGGAGAC	540
Db	161	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	180
QY	541	AGCCCTCGCAGGTGCTCTGCTGACTCAAAAGAAGAGTGGCTCGCGGCGCAGTGCCTC	600
Db	181	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTCGGTGTGACAGCGGCCACTGCTATGGATGATGATGATGATGATGATG	660
Db	201	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu	220
QY	661	GTCAGGCTTGAGAGTATGACCTCGCGCGCTGGAGAGTGGAGTGGAGTGGAGTGGAGT	720
Db	221	ValArgLeuGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	240
QY	721	AAGGAGTCTTCGTCCACCCCACTACAGAAAGACACACCGACCAATGATGATGATGATG	780
Db	241	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu	260
QY	781	CTGACCTGCGCCAGCGCGCCCTCTCGCAGACCATAGTGCATGCTGCTGCTGCTGCTG	840
Db	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	280
QY	841	AGCGGCTTCGAGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTTCGTGACGGGCTG	900
Db	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACACACAGCAGCAG	960
Db	301	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAAGATTCGCGTGGTCCGCGCACAAATGATGTCAGCAGAGTGCATGAGCAACATGGT	1020
Db	321	IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGAACATGTGTGTGCGGGCATCTCTCGGGGACCGCGCAGGATGCCTCGGAGGGGACAGT	1080
Db	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GGGGGGCCATGGTTCCTCTCCAGCGCACCTGTTCTGTTGGTGGCGCTGGTGGAGTGG	1140
Db	361	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	380
QY	1141	GGTGAAGGCTGTGGGCTCTTCACAACTACGGGCTTTTACACCAAGTACGCGCTACCTC	1200
Db	381	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	400
QY	1201	GACTGGATTCATGGGCACATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1245
Db	401	AspTrpIleHisGlyHisIleIleArgAspLysGluAlaProGlnLys	415

RESULT 3

## US-09-667-570A-3

; Sequence 3, Application US/09667570A

; Patent No. 6436397

; GENERAL INFORMATION:

; APPLICANT: Baker, Jeffrey C

; APPLICANT: Carlson, Andrew D

; APPLICANT: Huang, Lihua

; APPLICANT: Shelliga, Theodore A

; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C

; FILE REFERENCE: X-11796A

; CURRENT APPLICATION NUMBER: US/09/667.570A

; CURRENT FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: 60/045,255

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 419

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-667-570A-3

## Alignment Scores:

Seq. No.:	3,34e-167	Length:	419
Score:	2298.00	Matches:	415
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.33%	Indels:	0
DB:	4	Gaps:	0

US-09-763-153-1 (1-1245) x US-09-667-570A-3 (1-419)

QY	1	GCACACTCTCTCTGGAGGAGTCCGTCACAGACGCTGGAGCGGAGTGCATAGAGGAG	60
DB	1	AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	20
QY	61	ATCTGTGACTCGAGAGGCCAAGAAATTTCCAAATGTTGGATGACACACTGGCCCTC	120
DB	21	IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe	40
QY	121	TGGTCCAAAGCAGCTGCAGCGGTGACAGTGTGCTTGGCTTGGAGACCCGCTGGCC	180
DB	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	60
QY	181	AGCCTGTGCTGGCGCAGCGCAGTGCATCGACGCGCATCGCAGCTTCAGCTGCAGTGC	240
DB	61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
QY	241	CCAGCGGCTGGAGGGCGCTTCTGCCAGCGGAGTGCATTCCTCAATTGCTCGCTG	300
DB	81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
QY	301	GACACGGCGGCTGCACGATTAATCTAGAGAGTGGCTGAGAGTGGCTGAGTGTGT	360
DB	101	AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys	120
QY	361	GGCCTGTGCTACAGCTGGGGAGCAGCTTCTGCTGAGTGTCCACCGCAGTGAAGTCCCT	420
DB	121	AlaProGlyTyrLysLeuGlyAspSerLeuLeuGlnCysHisProAlaValLysPhePro	140
QY	421	TGTGGAGGCGCTGGAAGCGGTGAGAGAGAGCGCAGTCCAGTCAACAGCAGACAGAA	480
DB	141	CysGlyArgProTrpArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	160
QY	481	GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGAGAGATCACCAGCGGGAGAC	540
DB	161	AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgArgGlyAsp	180
QY	541	AGCCCTGTGCGAGTGGTCTCTGTGACTCAAGAGAGAGTGGCTGCGGGGCGAGTGC	600
DB	181	SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTCTGGTGTGTGACACGGCGCCACTGCATGGATGAGTCCAAAGCTCTT	560

DB	201	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	220
QY	661	GTCAGGCTGGAGAGTATGACCTGGCGGCTGGGAGAAAGTGGAGCTGGACCTGGACATC	720
DB	221	ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	240
QY	721	AAGGAGTCTCTCTCACCCCAACTACAGCAAGAGCACCACCAATGACATCGGACATG	780
DB	241	LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu	260
QY	781	CTCACCTGGCCCGCCGCCACCTCTCGCAGACCATAGTCCCATCTGCCCTCCGGAC	840
DB	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLecCysLeuProAsp	280
QY	841	AGCGGCTTGCAGAGCGCGAGCTCAATCAGCGCGGCGGAGGACCCCTCGTGGCGCTCG	900
DB	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACACAGCAGCGGAGAGGAGGAGCAAGAGAAACCGCACCTTCCTCTCAACTTC	960
DB	301	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAAGATTCCTGTCGCGCACATGAGTGCAGCGAGTGCATGAGCAACATGCTCT	1020
DB	321	IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGAATGCTGTGCGGCGCATCTCGGCGGCGGAGGATCGCTCGGAGGCGGACACT	1080
DB	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GGGGGGCCCATGTCCTCTTCCACGGCACCTGTTCTCTGTTGGGCTGTGAGCTGG	1140
DB	361	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	380
QY	1141	GTTGAGGCTGTGGCTCTTCCAACTACGCGGTTCACCAAGTTCAGCGCTACCTC	1200
DB	381	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValThrLysValSerArgTyrLeu	400
QY	1201	GACTGATCATCGGCGCATCATCAGAGCAAGAAAGCCCCCAGAG	1245
DB	401	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	415

## RESULT 4

PCT-US92-10242-1

; Sequence 1, Application PC/TUS9210242

; GENERAL INFORMATION:

; APPLICANT: Griffin, John H.

; APPLICANT: Mesters, Rolf

; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and

; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods

; TITLE OF INVENTION: for Inhibiting Coagulation

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Office of Patent Counsel, The Scripps

; ADDRESSEE: Research Institute

; STREET: 10666 North Torrey Pines Road, TPC 8

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; APPLICATION NUMBER: PCT/US92/10242

; FILING DATE: 19921118

; CLASSIFICATION:

; PRIOR APPLICATION NUMBER: US 07/793,989

; FILING DATE: 18-NOV-1991



## ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCRO472P  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Region

LOCATION: 1..157

OTHER INFORMATION: /note= "Protein C Light Chain"

FEATURE:

NAME/KEY: Region

LOCATION: 158..169

OTHER INFORMATION: /note= "Protein C Activation

OTHER INFORMATION: Peptide"

FEATURE:

NAME/KEY: Region

LOCATION: 170...419

OTHER INFORMATION: /note= "Protein C Heavy Chain"

PCT-US92-10242-1

## Alignment Scores:

Pred. No.: 3 34e-167 Length: 419  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 5 Gaps: 0

US-09-763-153-1 (1-1245) x PCT-US92-10242-1 (1-419)

QY 1 GCAACTCTCTCTGAGAGCTCTCTACAGCAGCGCTGGAGCGGAGTGCATAGAGGAG 60  
Db 1 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
QY 61 ATCTGTACTTCGAGGCGCCAGGAAATTTCCAAATGTGATGACACACTGGCCTTC 120  
Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40  
QY 121 TGTCTCAAGCACGTCGACGCTGACCACTGCTTGGTCTTGGCTTGGAGCACCCTGCGCC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCTGTGCTGCGGCGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGCTGGAGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAGCGGCTGTCACGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 101 AspAsnGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGGAGGCGCTGGAAGCGGATGGAAGAGAGCGGAGTGCATGAGTGAACAGACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAGACCAAGTAGATCGCGGCTCATTTGATGGGAAGATGACACCGGGGAGAC 540

Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp 180  
QY 541 AGCCCTCGGAGGTGGTCTCTGCTGACTCAAGAAGAGCTGGCTGGCGGGAGTGCCTC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTCTGGTGTCTGACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 220  
QY 661 GTCAGGCTTGGAGAGTATGACTGCGCGCTGGAGAGAGTGGAGCTGGAGCTGGACATC 720  
Db 221 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 260  
QY 781 CTGCACCTGGCGCCAGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
QY 841 AGCGGCTTGGAGAGCGGAGCTCAATCAGCGCCCGCCAGAGACCCCTGCTGCTGCTGCTG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGGAG 960  
Db 301 GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGCTGTGTGCGGCGCATCTCTGCGGAGCGGAGAGAGAGAGAGAGAGAGAG 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTCTCAAACTACCGCGCTTTACACCAAGTCCAGCGCTACCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpLysValSerArgTrpLeu 400  
QY 1201 GACTGGATCCATGGCGCATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 5

US-08-756-506-2

; Sequence 2, Application US/08756506

; Patent No. 5905185

; GENERAL INFORMATION:

; APPLICANT: Garner, Ian

; APPLICANT: Cottingham, Ian R.

; APPLICANT: Temperley, Simon M.

; APPLICANT: Foster, Donald C.

; APPLICANT: Sprecher, Cindy A.

; APPLICANT: Prunkard, Donna E.

; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESS: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
;

```

Alignment Scores:		
Prod. No.:	3,423-167	Length:
Score:	2398.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	98.33%	Indels:
DB:	2	Gaps:
		460
		415

US-09-763-153-1 (1-1245) x US-08-756-506-2 (1-460)

QY	1	GCCAACTCTTCTCGAGGAGCTCCGTACAGCAGCCTGGAGCGGAGTGCATAGAGGAG	60
DB	43	AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	62
QY	61	ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATTTGTGATGCACACCTGGCCCTC	120
DB	63	IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
QY	121	TGTTCCAAGCAGCTCGACGGTGACCAAGTCTTGGTCTTGCCTTGGAGCACCCGTGCGCC	180
DB	83	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
QY	181	AGCTGTGCTCGCGGCACGCGCATCGACGGCATCGGAGCTTCAGCTGCACATGC	240
DB	103	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
QY	241	CGCAGCGGCTGGAGGGCGGCTTCTCCACGCGGAGTGCAGTCTCCTCAATTGCTCGCTG	300
DB	123	ArgSerGlyTrpGluArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
QY	301	GACAACGCGCGCTGCACGCATTACTCCCTAGAGGAGTGGCGTGGCGCTGTAGCTGT	360
DB	143	AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys	162
QY	361	GCGCCTGGCTACAAGCTGGGGAGACCTCTCTCAGTGTACCCCGCGAGTGAAGTTCCT	420
DB	163	AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro	182
QY	421	TGTGGAGGCCCTGGAAGCGGATGGAGAAGCGCAGTCACCTGAAACGACACACAGAA	480
DB	183	CysGlyArgProIrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	202
QY	481	GACCAGAAGACCAAGTAGATCCGCGGCTCATTTGATGGGAAGATCACACGGCGGGAGAC	540
DB	203	AspGlnAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	222
QY	541	AGCCCTGGCAGTGTCTCTGTGGACTCAAGAGAGAGCTGCCCTGGCGGGCAGTGTC	600
DB	223	SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu	242

CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sawislak, Deborah A  
 REGISTRATION NUMBER: 37,438  
 REFERENCE/DOCKET NUMBER: 95-28  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6672  
 TELEFAX: 206-442-6678  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 460 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 08-756-506-4

Alignment Scores:  
 Pred. No.: 3,42e-167 Length: 460  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 2 Gaps: 0

US-09-763-153-1 (1-1245) x US-08-756-506-4 (1-460)

```

QY 1 GCCAACTCCTCTCGAGAGCTCCGTACACAGCAGCCTGGAGCGGGAGTGATAGAGGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGluGlu 62
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAAATTTTCCAAATGTGGATGACACACTGGCCCTTC 120
Db 63 IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCAACAGCAGTCGAGGTGACAGCTGCTTGTGCTTGGCCCTTGGAGACCCGTCGCC 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCCCTGTCTGGGGGACGACGTCATCGACGCGCATCGGCGCTTCAGCTTCAGCTGCAGTC 240
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGCTGGAGGGCGGCTTCTGCCAGCGGAGGTGAGCTTCTCAATGTCTCGCTG 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAspCysSerLeu 142
QY 301 GACACGGCGGTGACACCATCTACTCCCTAGAGAGGTGGCGTGGCGCGCTGTAGCTGT 360
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162
QY 361 GCGCTGTGCTACAAGCTGGGGACGACCTCCTGAGTGTACCCGCGCAGTGAAGTTCCT 420
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGGCCCTGGAAGCGGATGAGAGGAGCGAGTCACTACCTCAACAGCAGACAGAA 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGAGATGACACGCGGGGAGAC 540
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222
QY 541 AGCCCTCGCAGGTGCTCTGCTGACCTCAAGAGAGCTGGCGTGGGGGCGAGTGCCTC 600
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242
QY 601 ATCCACCCCTCTGCGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGAT 660
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262
QY 661 GTCAGGCTTGGAGAGTATGACCTCGCGGCTGGGAGAGAGTGGGAGCTGGACCTGGACATC 720
Db 263 ValArgLeuGlyGlyLysLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282

```

```

QY 721 AAGAGGCTCTCGTCCACCCCACTACAGCAAGAGCACCACGCAATGATGATGCACTG 780
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302
QY 781 CTGACACTGGCCCGAGCGCCACCTCTCGCAGACCATAGTCCCATCTGCTCCCGGAC 840
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322
QY 841 AGCGGCTTCAGAGCGCGAGCTCAATCAGCGCGCCAGGAGAGACCGACCTTCGCTCAACTTC 900
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342
QY 901 GGCTACACAGCAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362
QY 961 ATCAAGATTCCTGGTCCGCGCAATGAGTGCAGCGAGGTGCATGAGCAACATGCTGCT 1020
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382
QY 1021 GAGAACTGTCTGTCGGGCATCTCTCGGGACCGGAGAGTGCCTCGGAGGGCGACAGT 1080
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402
QY 1081 GGGGGGCCATGTCGCTCTCCACGCGACCTGCTTCTGTTGGCGCTGGTGGAGCTGG 1140
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422
QY 1141 GGTGAGGCTGTGGCTCTTCACTACAGTACGGGCTTTACACCAAGTCACCGCTACCTC 1200
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442
QY 1201 GACTGTGATCCATGGCGCATCATCAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1245
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

```

# RESULT 7

5225537-2  
 ; Patent No. 5225537  
 ; APPLICANT: FOSTER, DONALD  
 ; TITLE OF INVENTION: METHODS FOR PRODUCING HYBRID  
 ; PHOSPHOLIPID-BINDING PROTEINS  
 ; NUMBER OF SEQUENCES: 14  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/459,082  
 ; FILING DATE: 29-DEC-1989  
 ; SEQ ID NO: 2:  
 ; LENGTH: 461  
 5225537-2

Alignment Scores:  
 Pred. No.: 3,42e-167 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 2 Gaps: 0

US-09-763-153-1 (1-1245) x 5225537-2 (1-461)

```

QY 1 GCCAACTCCTCTCGAGAGCTCCGTACACAGCAGCCTGGAGCGGGAGTGATAGAGGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGluGlu 62
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAAATTTTCCAAATGTGGATGACACACTGGCGCTTC 120
Db 63 IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCAACAGCAGTCGAGGTGACAGCTGCTTGTGCTTGGCCCTTGGAGACCCGCTGCC 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102

```

QY 181 AGCTGTGCTCGGGACGGCACGCTGATCGACGGCATCGGACCTTCAGCTGCGACTGC 240  
Db 103 SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGGCGCTCTCCAGCGCGAGCTGAGCTTCCTCAATGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAAAGCGCGCTGACCAATCTACTGCTAGAGAGGTGGCTGGCGCGCTGAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTGCTACAGCTGGGGACGACCTCTGCTGAGTGTACCCCGGAGTTCCT 420  
Db 163 AlaProGlyTrpLysLeuGlyAspLeuGlyHisProAlaValLysPhePro 182  
QY 421 TGTGGAGCGCTGGAACGGATGGAGAGAGCGCAGTCACTGAAACGAGACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGAAAGATGACCGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGCAGGTGGTCTGCTGCTGACTCAAGAGAGAGTGGCTGCGGGCGAGTCTC 600  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
QY 661 GTCAGGCTTGGAGATGATGACTGCGCGCTGGGAGAGTGGAGCTGGAGCTGGACATC 720  
Db 263 ValArgLeuGlyGlyTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGAGGTCTTGGTCCACCCCACTACAGCAAGAGACACCGCACATGATGCTGCTGCT 780  
Db 283 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACCTTGGCGCCGCGCCGCTCTGCGAGACCATGAGTGGCTGCTGCTGCTGCTGCT 840  
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGCTTGGAGCGCGAGCTCAATCAGCGCGCGCCAGGAGACCCCTGCTGACGGGCTGG 900  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GCTACACAGCGCGAG 960  
Db 343 GlyThrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCCGTCCCGCACCAATCAGTGCAGCGAGGTGATGAGCAACATGGTGCT 1020  
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGAATCTGCTGTGCGGGCATCTCCGCGGACCGCGCAGGATGCTGCGAGGGCGACAGT 1080  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGGTGGCTCTCCACCGCACCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGGTGTGGGCTCTCTTCAACACTACGCGCTTTACACAAAGTCAAGCTGCTGCT 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 442  
QY 1201 GACTGGATCCATGGGCACATGAGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 8

5460953-3  
; Patent No. 5460953  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; GLYCOSYLATION MUTANTS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/93,217  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 628,063  
; FILING DATE: 21-DEC-1990  
; APPLICATION NUMBER: 484,081  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO: 3  
; LENGTH: 461  
5460953-3

Alignment Scores:  
Pred. No.: 9,78e-167 Length: 461  
Score: 2292.00 Matches: 414  
Percent Similarity: 99.76% Conservative: 0  
Best Local Similarity: 99.76% Mismatches: 1  
Query Match: 98.07% Indels: 0  
DB: 6 Gaps: 0

US-09-763-153-1 (1-1245) x 5460953-3 (1-461)

QY 1 GCCAATCTCTCTCTGAGGAGCTCGCTCAGCAGCTGGAGCGGAGTGCATAGAGGAG 60  
Db 43 AlaAsnSerPheLeuGluLeuArgHisSerLeuGluArgGluGlyCysIleGluGlu 62  
QY 61 ATCTGTGACTTTCGAGGAGCGCAAGGAAATTTCCAAAATGTGGATGACACTGGCCTTC 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGTCTCAAGACACTGACCGGTGACAGTGTGGTCTTGCCTTGGAGCACCGCTGCGCC 180  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCTGTGCTGGCGGACGCGCAGCTGCATCGCGCATCGGAGCTTCAGCTGCGACTGC 240  
Db 103 SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGGCGCGCTTCTGCCACGCGAGGTGAGCTTCTCAATGCTGCTGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAACGGCGGCTGACCGCATTTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGGCTTACAACTGGGGAGCGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 163 AlaProGlyTrpLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGCGCTTGGAGCGGATGGAGAGACCGCAGTCACTGAAACGAGACAGACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGACCAACTAGATCCGCGCTCATTTGATGGAAAGATGACCGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGCAGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
QY 661 GTCAGGCTTGGAGATGATGACTGCGCGCTGGGAGAGTGGAGCTGGAGCTGGACATC 720

Alignment Scores:					
Pred.	No.:	9.29e-165	Length:	461	
Score:	:	2266.00	Matches:	409	
Percent Similarity:	:	99.28%	Conservative:	3	
Best Local Similarity:	:	98.55%	Mismatches:	3	
Query Match:	:	96.96%	Indels:	0	
DB:	:	6	Gaps:	0	
US-09-763-153-1 (1-1245) x 5270178-17 (1-461)					
QY	1	GCCAACTCTTCTCGAGGAGTCGTGCACAGCAGCTGGAGCGGAGTGCATAGGAG	60		
Db	43	AlaAsnSerPheLeuGlulLeuArgHisSerSerLeuGluarGlucysIleGlulGlu	62		
QY	61	ATCTGTGACTTCGAGAGGCCAAGAATAATTTCCAAATGTGGATGACACACTGGCCTTC	120		
Db	63	IleCysAspPheGlulAlalLysGlulilePheGlnAsnValAspThrLeuAlaPhe	82		
QY	121	TGGTCCAAGCAGCTCGACGGTGACCAGTGTGCTCTTGCCCTGGAGCACCCGTGGCC	180		

## RESULT 10

5270178-18  
; Patent No. 5270178  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; ZYMOGEN FORMS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 21  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/484,133  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO:18:  
; LENGTH: 461  
5270178-18

Alignment Scores:  
Pred. No.: 9,29e-165 Length: 461  
Score: 2266.00 Matches: 409  
Percent Similarity: 99.28% Conservative: 3  
Best Local Similarity: 98.55% Mismatches: 3  
Query Match: 96.96% Indels: 0  
Gaps: 0

09-763-153-1 (1-1245) x 5270178-18 (1-461)

QY 1 GCCAACTCTTCTGAGGAGCTCCGTCACAGCAGCCCTGGAGCGGAGTGCATAGAGGAG 60  
|||||  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAATTTTCCAAATGTGGATGCACACATGGCCCTC 120  
|||||  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAACAGCTGACGAGTGCACGAGTGTGGTCTTGGCTTGGAGCACCCTGGCCGC 180  
|||||  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGGGGCAGCGCAGCTGCATCGACGCGCAGTCCGAGCTTCAGCTGGCAGTGC 240  
|||||  
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGGCTGGAGGGCGCTCTGCCAGCGGAGGTGAGTCTCAATTCGTCGGTG 300  
|||||  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAACGGGGCTGCACGATTAATCTACCTAGAGAGGTGGGTGGCGGCTGTAGCTGT 360  
|||||  
Db 143 AspAsnGlyGlyCysThrHisTyrcysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCCTGTGCTACAGCTGGGGCAGCGACCTCTCGAGTGTCCCGCAGTGAAGTTCCT 420  
|||||  
Db 163 AlaProGlyTyrllysLeuGlyAspAspLeuLeuGlnCysHisProAlaVallysPhePro 182  
QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGAGTCACTGAACAGCAGACAGAA 480  
|||||  
Db 183 CysGlyArgProTrpIysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACAGCGGGGAGAC 540  
|||||  
Db 203 AspGlnGluAspGlnValPheProArgLeuIleAsnGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGGCAGGTGCTGCTGGACTCAAGAAAGAGCTGCCCTGGCGGCGAGTGC 600  
|||||  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGGGTGTGACAGCGGCGGCTGATGATGATGATGATGATGATGATGAT 660  
|||||  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuIle 262  
QY 661 GTCAGGCTTCGAGGATGATACCTGGCGGCTGGGAGAGTGGAGTGGAGCTGGAGCATC 720  
|||||  
Db 263 ValArgIleGlyGluTyrlAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282

QY 721 AAGGAGTCTTCTGCTCCACCCCACTACAGCAAGAGACACACCGACAATGACATCGACTG 780  
|||||  
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACTGGCCAGCCCGCCACCCTCTCGCAGACCATAGTCCCATCTGCTCCCGGAC 840  
|||||  
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTTCAGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTCGCTGACGGCTG 900  
|||||  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACACAGCAGCCGCGAGAGAGAGGAGCAAGAGAAACCGCACTTCGCTCACTTC 960  
|||||  
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAGATTCCCGTGTCTCCCGCAATGAGTGCAGCGAGGTGCATGAGCAACATGGTGTCT 1020  
|||||  
Db 363 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGAAATGCTGTGTGGCGCATCTCTCGGGGACCGGCGAGGATCGCTCGAGGCGCAGT 1080  
|||||  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGGCCATGTGCTCTCTCCAGCGCACCTGTTCTGCTGGCGCTGGTGCAGCTGG 1140  
|||||  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGCTCTTCAACTACGCGGCTTTACACCAAGTTCAGCCGCTACCTC 1200  
|||||  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 442  
QY 1201 GACTGTGATTCATGGCGCATCAGAGCAAGAGAGGAGCGGCGGCGGCGGAG 1245  
|||||  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## RESULT 11

5270178-13  
; Patent No. 5270178  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; ZYMOGEN FORMS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 21  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/484,133  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO:13:  
; LENGTH: 460  
5270178-13

Alignment Scores:  
Pred. No.: 2,43e-164 Length: 460  
Score: 2260.50 Matches: 409  
Percent Similarity: 99.28% Conservative: 3  
Best Local Similarity: 98.55% Mismatches: 2  
Query Match: 96.73% Indels: 1  
Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-13 (1-460)

QY 1 GCCAACTCTTCTGAGGAGCTCCGTCACAGCAGCCCTGGAGCGGAGTGCATAGAGAG 60  
|||||  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAATTTTCCAAATGTGGATGCACACATGGCCCTC 120  
|||||  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAACAGCTGACGAGTGCACGAGTGTGGTCTTGGCTTGGAGCACCCTGGCGCC 180  
|||||  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGGGGCAGCGCAGCTGCATCGACGCGCAGTTCAGCTTCAGCTGGCAGTGC 240



```
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CCACGGCTGGAGGGCGCTTCGACGGCGAGGTGAGCTCCTCAATTGCTCGCTG 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACAACGGCGCTGCACCATTTACGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360
Db 143 AspAsnGlyGlyCysThrHisGlyCysLeuGluValGlyTrpArgCysSerCys 162
QY 361 GCGCTGGCTACAACTGGGAGACACTCCTGCAGTGTCAACCGGAGTGAAGTTCCT 420
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGGCGCTGGAACCGATGGAGAGAGCGCAGTCACTCAACGAGACACAGAA 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACGCGGGAGAC 540
Db 203 AspGlnGluAspGlnValAspProArgLeu---AsnGlyLysMetThrArgGlyAsp 221
QY 541 AGCCCTGGCAGGTGCTCCTGCTGACTCAAAAGAGAGTGGCTCGCGGCGAGTGTCT 600
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241
QY 601 ATCCACCCCTCCTGGGTGTGACAGCGCCACTGCATGATGATGATGATGATGATGAT 660
Db 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 261
QY 661 CTCAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 262 ValArgIleGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 281
QY 721 AAGAGGTCTTCGCCACCCCACTACAGCAAGAGCACCACCAATGATGATGATGATGAT 780
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301
QY 781 CTGACCTGGCGCCAGCCGCTCCTCGCAGACCATAGTGGCCATCTGCTGACGGCTGG 360
Db 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysLeuProAsp 321
QY 841 AGCGGCTTCGAGAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCTGTCAGCGGCTGG 900
Db 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyClnGluThrLeuValThrGlyTrp 341
QY 901 GGCTACACAGCAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 342 GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361
QY 961 ATCAAGATTCCTGCTGCGGCGCATCTCGGAGCAGGAGTCAATGAGCAACATGCTGCT 1020
Db 362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381
QY 1021 GAGAACATGCTGTGCGGCGCATCTCGGAGCAGGAGTCAATGAGCAACATGCTGCT 1080
Db 382 GluAsnMetLeuLysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401
QY 1081 GGGGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421
QY 1141 GGTGAGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrHisValSerArgTyrLeu 441
QY 1201 GACTGATCATGGGCACATCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245
Db 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456
RESULT 12
5270178-14
```

```
; Patent No. 5270178
; APPLICANT: GRINNELL, BRUCE E.; GRINNELL, BRIAN W.
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF
; ZMOGEN FORMS OF HUMAN PROTEIN C
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/484,133
; FILING DATE: 23-FEB-1990
; SEQ ID NO: 14:
; LENGTH: 460
5270178-14

Alignment Scores:
Pred. No.: 2,43e-164 Length: 460
Score: 2260.50 Matches: 409
Percent Similarity: 99.28% Conservative: 3
Best Local Similarity: 98.55% Mismatches: 2
Query Match: 96.73% Indels: 1
DB: Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-14 (1-460)
```

```
QY 1 GCCAACTCCTCTCTGGAGGAGCTCCCTCACAGCAGCTGGAGCGGAGTGCATAGAGAG 60
Db 43 AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
QY 61 ATCTGTGACTTCGAGGAGGCAAGAAATTTTCCAAATATGTGGATGACACACTGGCTTC 120
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCCAAACAGCAGTCACGCTGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCCTGTGCTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 103 SerLeuLysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACAAGCGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 143 AspAsnGlyGlyCysThrHisGlyThrCysLeuGluValGlyTrpArgArgCysSerCys 162
QY 361 GCGCTGCTGCTCAACGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGGCGCTGGAAGCGGATGAGTCCGCGCTCATTTGATGGAAGAGATGACACGAG 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAACCAAGTAGATCCGCGGCTCATTTGATGGAAGAGATGACACGAGGAGGAG 540
Db 203 AspGlnGluAspGlnValAspProArgLeu---AsnGlyLysMetThrArgArgGlyAsp 221
QY 541 AGCCCTGGCAGGTGCTCCTGCTGACTCAAAAGAGAGTGGCTCGCGGCGAGTGTCTGCT 600
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241
QY 601 ATCCACCCCTCCTGGGTGCTGACAGCGCCCACTGCATGATGATGATGATGATGATGAT 660
Db 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 261
QY 661 GTCAGGCTTGAGAGTATGACTCGCGGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 262 ValArgIleGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 281
QY 721 AAGAGGTCTTCGTCCACCCCACTACAGCAAGAGCACCACCAATGATGATGATGATGAT 780
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301
```

```
QY 781 CTGCACCTGGCCAGCCGCGCCACCTCTCGGACGACATAGTGCCTCTCCCTCCGGAC 840
Db 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321
QY 841 AGCGGCTTCGAGAGCGGAGCTCAATCAGCGCGCCAGGAGACCCCTCGTGACGGGTGG 900
Db 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341
QY 901 GGCTACACAGACGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 342 GlyTyrHisSerSerArgGluGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361
QY 961 ATCAAGATCCCGTGGTCCCGACCAATGAGTGCAGCGAGGTCTATGAGCAACATGGTCTCT 1020
Db 362 IleYsIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381
QY 1021 GAGAACATCTGTGTGGGGGATCTCTCGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401
1081 GGGGGGCCATGGTCCCTCTCCACGGCACCTGGTTCCTGGTGGGCTGGTGGTGGTGGTGGTGG 1140
402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421
QY 1141 GGTGAGGCTGTGGGCTCTCTTACAACTACGCGGCTTTACACCAAGATGACCGCTACCTC 1200
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 441
QY 1201 GACTGGATCCATGGGCATCATGAGACAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245
Db 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456

RESULT 13
US-09-065-872-1
; Sequence 1, Application US/09065872
; Patent No. 6162629
; GENERAL INFORMATION:
; APPLICANT: Baker, Jeffrey C
; APPLICANT: Carlson, Andrew D
; APPLICANT: Huang, Lihua
; APPLICANT: Shelliga, Theodore A
; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C
; FILE REFERENCE: apc process patent
; CURRENT APPLICATION NUMBER: US/09/065,872
; CURRENT FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: 60/045,255
; EARLIER FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
; 9-065-872-1

Alignment Scores:
Pred. No.: 6,2e-164 Length: 410
Score: 2255.00 Matches: 406
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.49% Indels: 0
DB: Gaps: 4

US-09-763-153-1 (1-1245) x US-09-065-872-1 (1-410)
QY 28 CACAGAGCCTGGAGGGAGTGCATAGAGGAGATCTGTGACTTCGAGGAGGCCAAGAA 87
Db 1 HisSerSerLeuGluArgGluCysIleGluGluIleCysAspPheGluAlaLysGlu 20
QY 88 ATTTTCCTCAAAATGTGATGACACACACACACACACACACACACACACACACACACACACAC 147
Db 21 IlePheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspGlyAspGln 40
```

```
QY 148 TGCCTTGTCTTCCCTTGGAGCACCCGCGCCAGCCTGTGCTGGGCGACGCGACGCTGC 207
Db 41 CysLeuValLeuProLeuGluHisProCysAlaSerLeuCysCysGlyHisGlyThrCys 60
QY 208 ATCGAGCGCATCGGAGCTTCAGCTGCGACTCCGCGAGCGGCTGGAGGGCGGCTCTTCTGC 267
Db 61 IleAspGlyIleGlySerPheSerCysAspCysArgSerGlyTrpGluGlyArgPheCys 80
QY 268 CAGCGGAGGTGAGCTTCCTCAATTCGCTGGAGCAACGCGGGCTGCACGCTACTACTGC 327
Db 81 GlnArgGluValSerPheLeuAsnCysSerLeuAspAsnGlyGlyCysThrHisTyrCys 100
QY 328 CTAGAGAGGTGGGCTGGCGGCTGTAGCTGTGCGCTGGCTACAACTGAGGTGGGGGACGAC 387
Db 101 LeuGluGluValGlyTrpArgArgCysSerCysAlaProGlyTyrLysLeuGlyAspAsp 120
QY 388 CTCTCTGAGTGTACCCCGCAGTGAAGTTCCTTGGGAGGCGCTTGGAAAGCGGTGGAG 447
Db 121 LeuLeuGlnCysHisProAlaValLysPheProCysGlyArgProTrpLysArgMetGlu 140
QY 448 AAGAAGCGAGTCACCTGAAACGAGACACAGAGAACCAAGAACCAAGATAGATCCGCGG 507
Db 141 LysLysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArg 160
QY 508 CTCAATGATGGGAGATGACCGCGGAGAGAGCGGCTGCATCCACCCCTCCCTGGTGGTGCAGCG 567
Db 161 LeuIleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValValLeuLeuAsp 180
QY 568 TCAAGAAGAAGCTGCGCGGCGAGTCTCATCCACCCCTCCCTGGTGGTGGTGCAGCG 627
Db 181 SerLysLysLysLeuAlaCysGlyAlaValLeuIleHisProSerTrpValLeuThrAla 200
QY 628 GCCACTGTCATGGATGAGTCCAAAGAGCTCTTGTGAGGCTTGGAGAGTATGACCTGCGG 687
Db 201 AlaHisCysMetAspGluSerLysLysLeuLeuValArgLeuGlyGlyTrpAspLeuArg 220
QY 688 CGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGAGAGTCTTCTCCACCCCACTAC 747
Db 221 ArgTrpGluLysTrpGluLeuAspLeuAspIleGlyGluValPheValHisProAsnTyr 240
QY 748 AGCAAGAGCACCGACCAATGACATGCGACTCTGCACCTGCGCCCGCCGCGCCACCCCTC 807
Db 241 SerLysSerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaThrLeu 260
QY 808 TCGCAGACCATAGTGCCCATCTGCCTCCCGGAGAGCGGCTTGCAGAGCGCGAGCTCAAT 867
Db 261 SerGlnThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuAsn 280
QY 868 CAGCGCGCGGAGAGACCCCTCGTGACGGGCTGGGCTTACCACAGACGCGGAGAGAGGAG 927
Db 281 GlnAlaGlyGlnGluThrLeuValThrGlyTrpGlyTyrHisSerSerArgGluLysGlu 300
QY 928 GCCAAGAGAAACCGCACCTTCTGCTCAACTTCATCAAGATTCCCGTGGTCCCGCACAAAT 987
Db 301 AlaLysArgAsnArgThrPheValLeuAsnPheIleLysIleProValValProHisAsn 320
QY 988 GAGTGCAGCGAGGTATGAGCAACATGCTGTCTGAGACATGCTGTGCGGCGCATCCTC 1047
Db 321 GluCysSerGluValMetSerAsnMetValSerGluAsnMetLeuCysAlaGlyIleLeu 340
QY 1048 GGGGACCGGAGGATGCTGCGAGGCGGACAGTGGGGGGCCCATGCTGCTGCTCTTCCAC 1107
Db 341 GlyAspArgGlnAspAlaCysGluLysSerGlyGlyProMetValAlaSerPheHis 360
QY 1108 GGCACCTGTTCTCGTGGGCTGTGAGCTGGGCTGAGGCTGTGGGCTGTGGCTTCTTCAAC 1167
Db 361 GlyThrTrpPheLeuValGlyLeuValSerTrpGlyGluGlyCysGlyLeuLeuHisAsn 380
QY 1168 TACGGCGTTTACACAAAGTACGCGCTACCTCCACTGGATCCATGGCAGCATCAGAGAC 1227
Db 381 TyrGlyValTyrThrLysValSerArgTyrLeuAspTrpIleHisGlyHisIleArgAsp 400
```

[illegible]

Db	402	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	421
Qy	1141	GGTGAGGC---TGTGGGCTCTTCACAACTAGCGGCTTTACACCAAAAGTCAGCGCTAC	1119
Db	422	GlyGluGlyGluCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyr	441
Qy	1198	CTCAGCTGGATCCATGGGCACATCAGACACAAGGAAGCCCCCCCAGAA	1245
Db	442	LeuAspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 16			
US-09-065-872-2			
; Sequence 2, Application US/09065872			
; Patent No. 6162629			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Jeffrey C			
; APPLICANT: Carlson, Andrew D			
; APPLICANT: Huang, Lihua			
; APPLICANT: Sheliga, Theodore A			
; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C			
; FILE REFERENCE: apc process patent			
; CURRENT APPLICATION NUMBER: US/09/065,872			
; CURRENT FILING DATE: 1998-04-24			
; EARLIER APPLICATION NUMBER: 60/045,255			
; EARLIER FILING DATE: 1997-04-28			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 409			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-065-872-2			
Alignment Scores:			
Pred. No.:		2,52e-163	Length: 409
Score:		2247.00	Matches: 405
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		96.15%	Indels: 0
DB:		4	Gaps: 0
US-09-763-153-1 (1-1245) x US-09-065-872-2 (1-409)			
Qy	31	AGCAGCTGGAGCGGGAGTGCATAGAGGAGATCTGTGACTTCGAGGAGGCCAAGAAATT	90
Db	1	SerSerLeuGluArgGluCysIleGluGluIleCysAspPheGluAlaLysGluIle	20
Qy	91	TTCCAAATGTGATGACACACTGGCCTTCTGGTCCAACACAGTCGACGGTGACCAATGC	150
Db	21	PheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspGlyAspGlnCys	40
Qy	151	TTGTGCTTGCCTTGAGCACCCGTCGCCAGCTGCTGCTGGCGGACGCGACGCGATGCATC	210
Db	41	LeuValLeuProLeuGluHisProCysAlaSerLeuCysGlyHisGlyThrCysIle	60
Qy	211	GACGGCATCGCAGCTTCAGCTGCGACTGCCGCGAGCGGCTGGGAGGGCGCTTCTGCCAG	270
Db	61	AspGlyIleGlySerPheSerCysAspCysArgSerGlyTrpGluGlyArgPheCysGln	80
Qy	271	CGCGAGGTGAGCTTCCTCAATTGCTCGCTGGACACCGCGGCTGCAACGCATTACTGCCTA	330
Db	81	ArgGluValSerPheLeuAsnCysSerLeuAspAsnGlyGlyCysThrHisTyrCysLeu	100
Qy	331	GAGGAGTGGGCTGGCGGCGCTGACTGTGGCTGGCTACAAAGCTGGGGACACCTC	390
Db	101	GluGluValGlyTrpArgArgCysSerCysAlaProGlyTyrLysLeuGlyAspAspLeu	120
Qy	391	CTGCAGTGTCAACCCGAGTGAAGTTCCTGTGGGAGGCCCTGGGAAGCGGATGAGAAG	450
Db	121	LeuGlnCysHisProAlaValLysPheProCysGlyArgProTrpLysArgMetGluLys	140
Qy	451	AAGCGCAGTCACTGAAACAGACACAGAAAGACCAAGAACCAAGTAGATCCGCGGCTC	510

Db 141 LysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArgLeu 160  
QY 511 ATTGATGGGAAGATGACACAGCGGGGAGAGACACCCCTGGCAGGTGGTCTGCTGACTCA 570  
Db 161 IleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValValLeuLeuAspSer 180  
QY 571 AAGAAGAAGCTGGCTGCGGGGAGTGTCTCATCCACCCCTCTGGTGTGTGACAGCGGC 630  
Db 181 LysLysLysLeuAlaCysGlyAlaValLeuIleHisProSerTrpValLeuThrAlaAla 200  
QY 631 CACTGCATGGATGCTCAAGAGCTCTGTCTGAGGCTGGAGATGATGACTGCGGGC 690  
Db 201 HisCysMetAspGluSerLysLysLeuValArgLeuGlyGluTrpAspLeuArgArg 220  
QY 691 TGGGAGAAGCTGGAGCTGACCTGACATCAGAGGAGTCTGCTCCACCCCACTACAGC 750  
Db 221 TrpGluLysTrpGluLeuAspLeuAspIleLysGluValPheValHisProAsnTrpSer 240  
QY 751 AAGAGCACACCGACAATGACATCGACTGCTGACCTGGCCGAGCGCCACCCCTCTCG 810  
Db 241 LysSerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSer 260  
QY 811 CAGACCATAGTCCCATCTGCTCCGAGACAGCGGCTTGGAGCGGAGCTCAATCAG 870  
Db 261 GlnThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuAsnGln 280  
QY 871 GCCGGCAGGAGACCTCTGTCAGCGGCTGGGCTACACAGCAGCGGAGAGAGGAGCC 930  
Db 281 AlaGlyGlnGluThrLeuValThrGlyTrpGlyTrpHisSerArgGluLysGluAla 300  
QY 931 AAGAGAACCGCACCTTCTGCTCAACTTCATCAAGATTCCCGTGGTCCCGCACCAATGAG 990  
Db 301 LysArgAsnArgThrPheValLeuAsnPheIleLysIleProValValProHisAsnGlu 320  
QY 991 TGCAGGAGCTCATGAGCAACATGTGTCTGAGAACATCTGTGTGGCGGATCCCTCGGG 1050  
Db 321 CysSerGluValMetSerAsnMetValSerGluAsnMetLeuLysAlaGlyIleLeuGly 340  
QY 1051 GACCGCAGGAGTCTGCGAGGGGACAGTGGGGGCGCCATGGTCCCTCTCCACGGC 1110  
Db 341 AspArgGlnAspAlaCysGluLysAspSerGlyGlyProMetValAlaSerPheHisGly 360  
QY 1111 ACTCGTCTCTGGTGGGCTGTGAGCTGGGCTGAGGCTGTGGGCTCTTCCAACTAC 1170  
Db 361 ThrTrpPheLeuValGlyLeuValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTyr 380  
QY 1171 GCGCTTTACACCAAGTCAGCGCTTACCTCGACTGGATCCATGGCAGCATCAGAGACAG 1230  
Db 381 GlyValTyrThrLysValSerArgTyrLeuAspTrpIleHisGlyHisIleArgAspLys 400  
QY 1231 GAAGCCCCCAGAG 1245  
Db 401 GluAlaProGlnLys 405

## RESULT 17

US-09-667-570A-2  
; Sequence 2, Application US/09667570A  
; Patent No. 6436397  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Jeffrey C  
; APPLICANT: Carlson, Andrew D  
; APPLICANT: Huang, Lihua  
; APPLICANT: Sheliga, Theodore A  
; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C  
; FILE REFERENCE: X-11796A  
; CURRENT APPLICATION NUMBER: US/09/667,570A  
; PRIORITY FILING DATE: 2000-09-21  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 409

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-667-570A-2

## Alignment Scores:

Pred. No.: 2,52e-163 Length: 409  
Score: 2247.00 Matches: 405  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 96.15% Indels: 0  
DB: 4 Gaps: 0

US-09-763-153-1 (1-1245) x US-09-667-570A-2 (1-409)

QY 31 AGCAGCTGGAGCGGAGTGCATAGAGAGATCTGTGACTTCGAGGAGGCCAAGAAATT 90  
Db 1 SerSerLeuGluArgGluCysIleGluGluIleCysAspPheGluGluAlaLysGluIle 20  
QY 91 TTCCAAAATGTGGATGACACACTGGCCCTTGTGTCCTCAAGCAGCTCGAGCGGTGACCATGC 150  
Db 21 PheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspGlyAspGlnCys 40  
QY 151 TTGCTCTTGGCTTGGAGCACCCGTCGCCAGCTGTGCTCGCGGACAGCGCATTTACTGCATC 210  
Db 41 LeuValLeuProLeuGluHisProCysAlaSerLeuCysCysGlyHisGlyThrCysIle 60  
QY 211 GACGGCATCGGAGCTTACGCTGGAGTGGCGGAGCGGCTGGAGGCGCCCTTCTGCGCAG 270  
Db 61 AspGlyIleGlySerPheSerCysAspCysArgSerGlyTrpGluGlyArgPheCysGln 80  
QY 271 CGCAGGTGAGCTTCTCAATTGCTCGCTGACACAGCGCGCTGCGCATTTACTGCCTA 330  
Db 81 ArgGlnValSerPheLeuAsnCysSerLeuAspAsnGlyGlyCysThrHisTyrCysLeu 100  
QY 331 GAGAGGTGGGCTGGCGGCGCTGTAGCTGTGCGCTGCTACAGCTGGGGGAGCGACCTC 390  
Db 101 GluGluValGlyTrpArgArgCysSerCysAlaProGlyTyrLysLeuGlyAspAspLeu 120  
QY 391 CTGCAGTGTCAACCGGAGTGAAGTTCCTTGTGGAGGCGCTGGAGCGGATGGAGAG 450  
Db 121 LeuGlnCysHisProAlaValLysPheProCysGlyArgProTrpLysArgMetGluLys 140  
QY 451 AAGCGAGTCACTGAAACGAGACACAGAGACCAAGACCAAGTAGATTCGCGGCTC 510  
Db 141 LysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArgLeu 160  
QY 511 ATTGATGGGAAGATGACACAGCGGGGAGACAGCCCTTGGCAGGTGTCTGCTGGACTCA 570  
Db 161 IleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValLeuLeuAspSer 180  
QY 571 AAGAAGAAGCTGGCGGCGGAGTGTCTCATCCACCCCTCTGGTGTGACAGCGGCC 630  
Db 181 LysLysLysLeuAlaCysGlyAlaValLeuIleHisProSerTrpValLeuThrAlaAla 200  
QY 631 CACTGCATGGATGATCCAAAGAGCTCTTGTGAGGCTTGGAGAGTATGACTTCGCGCGC 690  
Db 201 HisCysMetAspGluSerLysLysLeuValArgLeuGlyGluTrpAspLeuArgArg 220  
QY 691 TGGGAGAAGTGGAGCTGGACCTGGAGCATCAAGAGGTCTTCCTCACCCCACTACAGC 750  
Db 221 TrpGluLysTrpGluLeuAspLeuAspIleLysGluValPheValHisProAsnTyrSer 240  
QY 751 AAGAGCACACCGACAATGACATCGACTGCTGACCTGGCCGAGCGCCACCCCTCTCG 810  
Db 241 LysSerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSer 260  
QY 811 CAGACCATAGTCCCATCTGCTCCGAGCAGCGGCTTGGAGCGGAGCTCAATCAG 870  
Db 261 GlnThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuAsnGln 280  
QY 871 GCCGGCAGGAGACCTCTGTCAGCGGCTGGGCTACACAGCAGCGGAGAGAGGAGCC 930  
Db 281 AlaGlyGlnGluThrLeuValThrGlyTrpGlyTyrHisSerArgGluLysGluAla 300

```
QY 931 AAGAGAAACCGACCTTCTGCTTCAACTTCATCAAGATTCCCGTGGTCCGACAAATGAG 990
|
|
|
Db 301 LysArgAsnArgThrPheValLeuAsnPheIleLysIleProValProHisAsnGlu 320
|
|
|
QY 991 TGCAGCGAGGTCTAGCAGACATGGTGTCTGAGAACATGTCTGTGTGGGGATCCTCGGG 1050
|
|
|
Db 321 CysSerGluValMetSerAsnMetValSerGluAsnMetLeuCysAlaGlyIleLeuGly 340
|
|
|
QY 1051 GACCGGACAGGATGCTGCGAGGCGCAGAGTGGGGGCCCATGGTGCCTCTTCCACGCG 1110
|
|
|
Db 341 AspArgGluAspAlaCysGluGlyAspSerGlyGlyPrometValAlaSerPheHisGly 360
|
|
|
QY 1111 ACCTGGTCTCTGGTGGGCTGGTGGTGGGCTGAGGCTGGGCTGCTTCAACATAC 1170
|
|
|
Db 361 ThrTrpPheLeuValGlyLeuValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTyr 380
|
|
|
QY 1171 GCGGTTTACAAAGTCAGCGCTACCTGACCTGAGTCCATGGATCCATGACATCAGACAG 1230
|
|
|
Db 381 GlyValTyrThrLysValSerArgTyrLeuAspTrpIleHisGlyHisIleArgAspLys 400
|
|
|
1231 GAAGCCCCCAGAG 1245
|
|
|
401 GluAlaProGlnLys 405
|
|
|
RESULT 18
5270178-15
; Patent No. 5270178
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF
; ZYMOGEN FORMS OF HUMAN PROTEIN C
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/484,133
; FILING DATE: 23-FEB-1990
; SEQ ID NO:15:
; LENGTH: 460
5270178-15

Alignment Scores:
Pred. No.: 3,37e-163 Length: 460
Score: 2245.50 Matches: 407
Percent Similarity: 98.80% Conservative: 3
Best Local Similarity: 98.07% Mismatches: 4
Query Match: 96.08% Indels: 1
Db: 6 Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-15 (1-460)

QY 1 GCCAACTCCTTCTGAGGAGCTCGTCCACAGACGCTTGAGCGGAGTGCATAGAGGAG 60
|
|
|
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
|
|
|
61 ATCTGTGACTCGAGGAGGCGCAAGAAATTTCCAAATGTGGATGCACACTGGCCTTC 120
|
|
|
63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82
|
|
|
QY 121 TGGTCCAAAGCAGCTCGACGGTGACCAAGTGTGTGTGCTTGGCCCTTGGAGCACCCGCTGGCGC 180
|
|
|
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
|
|
|
QY 181 AGCCTGTGCTGGGCGCAGCGCATGTCATGACGCGATCGGACGCTTCAGCTGGCAGTGC 240
|
|
|
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
|
|
|
QY 241 CCACGGCTGGAGGCGGCTTCTGCCAGCGCGAGGTGACTTCCCTCAATTGCTCGGTG 300
|
|
|
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
|
|
|
QY 301 GACACGCGCGCTGCACCATCTACTGCTCCTAGAGGAGTGGGCTGGCGCGCTGTAGCTGT 360
|
|
|
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162
|
|
|
```

```
QY 361 GCGCCTGGCTACAAGCTGGGGAGCGACCTCTCGAGTGTACCCCGCCAGTGAAGTTCCT 420
|
|
|
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182
|
|
|
QY 421 TGTGGGAGCGCTTGGAGCGGATGGAGAGAAAGCGCAGTCACCTGAAACGACACACAGAA 480
|
|
|
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
|
|
|
QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACCGGGGGAGAC 540
|
|
|
Db 203 AspGlnGluAspProValPheProArgLeu---AsnGlyLysMetThrArgArgGlyAsp 221
|
|
|
QY 541 AGCCCTTGGCAGGTGTCTGTGGACTCAAAGAGAACGCTGGCCTGGCGGCGAGTGCCTC 600
|
|
|
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 241
|
|
|
QY 601 ATCCACCCCTCTGGGTGTGTGACAGCGGCCCATGCTGATGATGATGATGATGATGATG 660
|
|
|
Db 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspLeuSerLysLysLeuIle 261
|
|
|
QY 661 GTCAGGCTTGGAGATGATGACCTGCGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720
|
|
|
Db 262 ValArgIleGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 281
|
|
|
QY 721 AAGGAGGTCTTGTCCACCCCACTACAGACAGAGCACCACGACATGACATCGGACTG 780
|
|
|
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 301
|
|
|
QY 781 CTCGACCTGGCCCGACCGCCCGCCCTCTCGACACCATAGTCCCATCTGCCCGGAC 840
|
|
|
Db 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321
|
|
|
QY 841 AGCGGCTTTCGAGCGCGAGCTCAATCAGCGCGCGGAGGAGACCCCTCGTGCAGGCGCTG 900
|
|
|
Db 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341
|
|
|
QY 901 GGCTACACAGCAGCGAGAGAGAGGAGGCGCAAGAGAAACCGCACCTCTGCTCTCACTTC 960
|
|
|
Db 342 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361
|
|
|
QY 961 ATCAAGATTCCTGCTGCCGACCAATGATGTCAGCGAGGTCTATGACCAATGCTGTCT 1020
|
|
|
Db 362 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381
|
|
|
QY 1021 GAGAACATGTGTGTGCGGCGCATCTCTCGGCGACCGCGGAGGAGGAGGCGGCGAGT 1080
|
|
|
Db 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401
|
|
|
QY 1081 GGGGGCGCATGTGCTCTCTTCCAGCGACCTGTTCTGCTGGGCGCTGTGAGCTGG 1140
|
|
|
Db 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421
|
|
|
QY 1141 GGTGAGGCTGTGGCTCTCTTCAAACTAGCGGCTTTTACACAAAGTACAGCCGCTACCTC 1200
|
|
|
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 441
|
|
|
QY 1201 GACTGTGATCCATGGGCACATCAGAGACAAAGAGAGGAGGAGGAGGAGGAGGAGG 1245
|
|
|
Db 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456
|
|
|
RESULT 19
5270178-16
; Patent No. 5270178
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF
; ZYMOGEN FORMS OF HUMAN PROTEIN C
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/484,133
; FILING DATE: 23-FEB-1990
; SEQ ID NO:16:
; LENGTH: 460
5270178-16
```



## Alignment Scores:

Pred. No.: 3.81e-161 Length: 460  
Score: 2218.50 Matches: 403  
Percent Similarity: 98.07% Conservative: 7  
Best Local Similarity: 97.11% Mismatches: 4  
Query Match: 94.93% Indels: 1  
DB: 6 Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-16 (1-460)

```
QY 1 GCCAAGTCTCTCTGGAGAGCTCCGTACAGCAGCTGGAGCGGAGTGCATAGGAG 60
43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 62
61 ATCTGTGACTTCGAGAGGCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCCTC 120
63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82
QY 121 TGGTCCAAAGCAGTCGACGCTGACAGTCTTGGTCTTGGCTTGGAGCACCCTGGCC 180
83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCCTGTCTGCGGGCAGCGCAGCTGCATCGACGCGCATCGGAGCTTCAGCTGCACCTGC 240
103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGCTGGAGGGCGCTCTCTGCCAGCGGAGGTAGCTTCTCAATTGCTCGGTG 300
123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACAAGCGGCTGACGACGATTAATCTCTAGAGGAGTGGCTGGCGGCTGTAGCTGT 360
143 AspAsnGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 162
QY 361 GCSCCTGGCTACAAGCTGGGGAGCGACCTCTCTGCACTGTCACCGCGAGTGAAGTTCCT 420
163 AlaProGlyTyrlsLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGGAGTCACTGAAAGAGACACAGAA 480
183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAGACCAAGTACATCGCGGCTCATTGATGGAGAGTACACAGCGGGGAGAC 540
203 LeuHisLysLeuGlnThrTyrlProArgLeu---AsnGlyLysMetThrArgGlyAsp 221
541 AGCCCTGGCAGGTGCTCTGTGGACTCAAGAGAGAGCTGGCTGGCGGCGAGTGTCTC 600
222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241
QY 601 ATCCACCCCTCTCGGTGTGTGACAGCGGCGCCTGCTGATGATGAGTCCAAAGAGCTCT 660
242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuIle 261
QY 661 GTGAGCTGGAGATGACCTGGCGGCTGGGAGAGTGGAGCTGGACCTGGGACATC 720
262 ValArgIleGlyLysTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 281
QY 721 AAGGAGGTCTCTGCCACCCCACTACAGCAGAGACACACCGACCAATCATCGACATG 780
282 LysGluValPheValHisProAsnTyrlSerLysSerThrThrAspAsnAspIleAlaLeu 301
QY 781 CTCACCTGGCCGCCGCCGCCCTCTCGCAGACATAGTCCCATCTGCCCTCCCGGAC 840
302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321
QY 841 AGCGGCTTTCAGAGCGCGAGCTCATATCAGCGCGGCGCAGGACACCTCTGTGAGCGGCTGG 900
322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341
QY 901 GGTCTACACAGCAGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
```

```
Db 342 GlyTyrlHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361
QY 961 ATCAAGATTCCTGGTCCCGCACAAATGAGTGCAGCGAGGTGATGACACATGGTGTCT 1020
Db 362 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381
QY 1021 GAGAACATGCTGTGTGGGGCATCTCTCGGGGAGCGGCGGAGGATGCTCGAGGGGACAGT 1080
Db 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401
QY 1081 GGGGGGCGCATGTGTGGCTCTCTCCAGCGCACCTGGTTCCTGGTGGCGCTGGTGGCTGG 1140
Db 402 GlyGlyPrometValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421
QY 1141 GGTGAGGCTGTGGGCTCTCTCACAACACTAGCGGCTTTACACCAAGTTCAGCCGCTACTC 1200
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValThrHisValSerArgTyrlLeu 441
QY 1201 GACTGGATCCATGGGCGACATCAGACAGCAAGGAGCGGCGGCGGAGG 1245
Db 442 AspIrlleHisGlyHisIleArgAspLysGluAlaProGlnLys 456
RESULT 20
US-07-720-189-1
; Sequence 1, Application US/07720189
; Patent No. 5279956
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf M.
; TITLE OF INVENTION: APC POLYPEPTIDES AND ANTI-PEPTIDE
; TITLE OF INVENTION: ANTIBODIES, DIAGNOSTIC METHODS AND SYSTEMS FOR INHIBITING
; TITLE OF INVENTION: APC, AND THERAPEUTIC METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: the Scripps Research Institute, Office of Patent
; ADDRESSEE: Counsel
; STREET: 3366 No. 5279956th Torrey Pines Court, Suite 240
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07720189
; FILING DATE: 19910724
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bingham, Douglas
; REGISTRATION NUMBER: 32,457
; REFERENCE/DOCKET NUMBER: SCRO390P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..262
; OTHER INFORMATION:
; /note= "In SEQ ID NO 1 is the sequence for
; the PC heavy chain, the amino acid residue
; positions of
; OTHER INFORMATION: which begin at position 158 and end at 419."
US-07-720-189-1
```

## Alignment Scores:

Pred. No.: 2.06e-98 Length: 262  
 Score: 1393.00 Matches: 258  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 59.61% Indels: 0  
 DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x US-07-720-189-1 (1-262)

```

QY 472 GACACAGAACCAAGACCAAGACCAAGTAGATCGCGGCTCATGTGGGAAGATGACCAGG 531
DB 1 AspThrGluAspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArg 20
QY 532 CGGGGAGACAGCCCTCGGCGAGTGTCTGCTGGACTCAAGAAGAAAGCTGGCCTGCGGG 591
DB 21 ArgGlyAspSerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGly 40
QY 592 GCAGTCTCATCCACCCCTCTCGGTGTGTGACAGGGCCCACTGCATGGATGAGTCCAAAG 651
DB 41 AlaValLeuIleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLys 60
652 AAGCTCTCTGTGAGGCTTGGAGAGTATGACCTGGGCGCTGGAGAGAGTGGGAGCTGGAC 711
DB 61 LysLeuLeuValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAsp 80
QY 712 CTGGACATCAAGAGGTCTCTCCACCCCACTACAGAACAGACACCCAGCAATGAC 771
DB 81 LeuAspIleLysGluValPheValHisProAsnTrpSerLysSerThrAspAsnAsp 100
QY 772 ATCGCACTGCTCCACCTGGCGCCAGCCCGCCACCTCTCGGACAGACCATAGTCCCATCTGC 831
DB 101 IleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCys 120
QY 832 CTCGCGGACAGCGGCTTTCAGAGCGGAGCTCAATCAGGCGCGGAGGAGACCCCTCGTG 891
DB 121 LeuProAspSerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuVal 140
QY 892 ACGGCTGGGGCTACACAGAGCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 951
DB 141 ThrGlyTrpGlyTyrosHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheVal 160
QY 952 CTCACACTTCACAGATTCCTCGGTGTCCTCCGACACATGAGTGCAGGAGGTCATGAGCAAC 1011
DB 161 LeuAsnPheIleLysIleProValValProHisAsnGluCysSerGluValMetSerAsn 180
QY 1012 ATGTTGTCTGAGACATCTGTGTGGGGCATCTCGGGGACCGGAGGATGCTGCGAG 1071
DB 181 MetValSerGluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGlu 200
QY 1072 GCGACAGTGGGGCCCATGTCGCTCTCTCCACGCGACCTGGTTCCTGGTGGGCTG 1131
DB 201 GlyAspSerGlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeu 220
1132 GTGAGCTGGGTGAGGGCTGTGGGCTCTCTCACAACTACGCGCTTTACACCAAGTCAGC 1191
DB 221 ValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTyThrLysValSer 240
QY 1192 CGCTACCTCGACTGGATCTGATGGACATCAGACAGCAAGGAGGAGGAGGAGGAGGAGGAG 1245
DB 241 ArgTyrLeuAspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 258
  
```

Search completed: March 17, 2003, 17:38:42  
 Job time : 56 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:35:25 ; Search time 47 Seconds  
(without alignments)  
2441.899 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

oring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 398832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09763153/runat\_11032003\_084249\_2120/app\_query.fasta\_1.1415  
-DB=PublishedApplications\_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=60 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09763153.acgn.1.1.12 @runat\_11032003\_084249\_2120  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pap.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pap.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2298	98.3	419	9	US-10-182-263-1
2	2298	98.3	419	9	US-09-978-917A-4
3	2298	98.3	461	9	US-10-182-263-2
4	2298	98.3	461	9	US-09-978-917A-2

#### ALIGNMENTS

US-10-182-263-1

; Sequence 1, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182.263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/181948

5	2270	97.1	419	9	US-10-182-263-5	Sequence 5, Appli
6	2264	96.9	419	9	US-10-182-263-3	Sequence 3, Appli
7	2262	96.8	419	9	US-10-182-263-6	Sequence 6, Appli
8	2260	96.7	419	9	US-10-182-263-4	Sequence 4, Appli
9	736	31.5	461	9	US-10-132-829-5	Sequence 5, Appli
10	736	31.5	461	10	US-09-884-901-3	Sequence 3, Appli
11	735	31.5	415	10	US-09-118-748-2	Sequence 2, Appli
12	726	31.1	406	10	US-10-109-498-1	Sequence 1, Appli
13	470	20.1	802	9	US-09-978-295A-169	Sequence 169, App
14	470	20.1	802	9	US-09-978-697-169	Sequence 169, App
15	470	20.1	802	9	US-09-978-192A-169	Sequence 169, App
16	470	20.1	802	9	US-09-999-832A-169	Sequence 169, App
17	470	20.1	802	9	US-09-978-189-169	Sequence 169, App
18	470	20.1	802	9	US-09-978-608A-169	Sequence 169, App
19	470	20.1	802	10	US-09-888-615-113	Sequence 113, App
20	428.5	18.3	850	10	US-09-888-615-108	Sequence 108, App
21	426.5	18.2	249	9	US-09-961-721-5	Sequence 5, Appli
22	426.5	18.2	251	9	US-09-874-198-6	Sequence 4, Appli
23	424.5	18.2	638	9	US-09-808-602-102	Sequence 102, App
24	420	18.0	679	10	US-09-874-238-6	Sequence 6, Appli
25	419.5	18.0	812	10	US-09-335-325-1	Sequence 1, Appli
26	419.5	18.0	812	9	US-10-131-241-1	Sequence 1, Appli
27	419.5	18.0	812	10	US-09-788-142-1	Sequence 1, Appli
28	419.5	18.0	812	10	US-09-761-120-1	Sequence 1, Appli
29	419.5	18.0	812	10	US-09-873-676-81	Sequence 81, Appli
30	419.5	18.0	625	9	US-09-808-602-104	Sequence 104, App
31	419	17.9	643	9	US-09-808-602-103	Sequence 103, App
32	417.5	17.9	855	10	US-09-900-751-2	Sequence 2, Appli
33	417.5	17.9	560	10	US-09-912-559-3	Sequence 3, Appli
34	410	17.5	791	9	US-09-967-386-1	Sequence 1, Appli
35	410	17.5	822	10	US-09-946-893-2	Sequence 2, Appli
36	410	17.5	822	10	US-09-147-947-6	Sequence 6, Appli
37	408.5	17.4	227	9	US-10-045-367A-4	Sequence 4, Appli
38	406.5	17.4	393	9	US-10-012-896-934	Sequence 934, App
39	405.5	17.4	393	9	US-09-895-793-934	Sequence 934, App
40	405.5	17.4	393	9	US-09-895-814-934	Sequence 934, App
41	405.5	17.4	393	10	US-09-759-143-934	Sequence 934, App
42	405.5	17.4	393	10	US-09-780-669-934	Sequence 934, App
43	405.5	17.4	393	10	US-09-822-827-934	Sequence 932, App
44	405.5	17.4	492	9	US-10-012-896-932	Sequence 932, App
45	405.5	17.4	492	9	US-09-895-793-932	Sequence 932, App
46	405.5	17.4	492	9	US-09-895-814-932	Sequence 932, App
47	405.5	17.4	492	10	US-09-759-143-932	Sequence 932, App
48	405.5	17.4	492	10	US-09-780-669-932	Sequence 932, App
49	405.5	17.4	492	10	US-09-822-827-932	Sequence 932, App
50	405.5	17.4	384	9	US-09-981-353-23	Sequence 23, Appli
51	402.5	17.2	492	9	US-10-012-896-895	Sequence 895, App
52	402.5	17.2	492	9	US-09-895-793-895	Sequence 895, App
53	402.5	17.2	492	9	US-09-895-814-895	Sequence 895, App
54	402.5	17.2	492	10	US-09-759-143-895	Sequence 895, App
55	402.5	17.2	492	10	US-09-780-669-895	Sequence 895, App
56	402.5	17.2	492	10	US-09-822-827-895	Sequence 14, Appli
57	402.5	17.2	492	10	US-09-822-827-895	Sequence 895, App
58	402.5	17.2	560	10	US-09-912-559-4	Sequence 4, Appli
59	402	17.2	855	9	US-10-099-700A-2	Sequence 2, Appli
60	401.5	17.2				

; PRIOR FILING DATE: 2002-02-11  
 ; PRIOR APPLICATION NUMBER: 60/189199  
 ; PRIOR FILING DATE: 2000-03-14  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 419  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-182-263-1

Alignment Scores:  
 Pred. No.: 1,48e-127 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-10-182-263-1 (1-419)

```

1  GCCAACTCCTCTCTGAGGAGCTCCGTACACAGCCTGGAGCGGAGTGCATAGAGGAG 60
1  AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 20
61  ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120
21  IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40
121  TGTCTCAAGCACATCGACGGTGACAGTGTGGTCTGGCTTCCCTTGGAGCACCGCTGGCC 180
41  TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60
181  AGCTGTGCTCGCGGACGCGACGTCATCGACGCGATCGGAGCTTCAGCTCGCACTGC 240
61  SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80
241  CGCAGCGGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGTTCCTCAATGTCTGCTG 300
81  ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100
301  GACACGCGGCTGACGAGTACTCGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTCT 360
101  AspAsnGlyGlyCysThrHisTrpCysLeuGluGluValGlyTrpArgCysSerCys 120
361  GCGCTGCTACAGTGGGGGACGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
121  AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140
421  TGTGGGAGGCGCTGGAGCGGATGGAGAGAACGCGAGTCACCTGAAACGAGACAGAA 480
141  CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160
481  GACCAAGACAGCAAGTAGATCGCGGCTCATTTGATGGGAAGATGACAGCGCGGAGAC 540
161  AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180
541  AGCCCTCGGAGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
181  SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200
601  ATCCACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
201  IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 220
661  GTCAGGCTTGGAGAGTATGACCTCGCGGCTGGGAGAGTGGAGCTGGAGCTGGAGCATC 720
221  ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240
721  AAGGAGGTCTTCGTCACCCCACTACAGCAAGAGCACCACGACCAATGATCGCACTG 780
241  LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260
  
```

```

781  CTGCACCTGGCCCGCCAGCCGCCCTCTCCAGACCATAGTGCCTATCTGCCTCCGGAC 840
261  LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280
841  AGCGGCTTCGAGAGCGCGAGCTCAATCAGCGCGCCAGGAGACCTCGTGACGGCTGG 900
281  SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300
901  GGCTTACCACAGCAGCGCCGAGAGAGGAGGCAAGAAACCCACCTTCCTCTCAACTTC 960
301  GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320
961  ATCAAGATTCCCGTGGTCCCGCACAATGAGTGCAGCAGGATCATGACCAACATGGTGTCT 1020
321  IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340
1021  GAGAACATGCTGTGCGGCGCATCTCGGAGACCGCGGAGGATGCTCGAGGCGGACAGT 1080
341  GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360
1081  GGGGGGCGCATGTCGCTCTCTCCACGCGACCTGCTTCCTGGTGGCGCTGGTGAGCTGG 1140
361  GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380
1141  GGTGAGGCTGTGGCTCTCAACTACGCGCTTTACACCAAGTCAGCGCTACCTC 1200
381  GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTyrLeu 400
1201  GACTGTGATCCATGGGCGCATCAGACAGCAAGGAAGCCCGCCAGAAAG 1245
401  AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415
  
```

# RESULT 2

US-09-763-153-1 (1-1245) x US-09-763-153-1 (1-419)  
 ; Sequence 4, Application US/09978917A  
 ; Publication No. US20030027299A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Maxygen Aps; Maxygen Holdings  
 ; TITLE OF INVENTION: Protein C or activated protein C-like molecules  
 ; FILE REFERENCE: 0219us310 - protein C  
 ; CURRENT APPLICATION NUMBER: US/09/978, 917A  
 ; CURRENT FILING DATE: 2001-10-17  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 419  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-763-153-1 (1-1245) x US-09-763-153-1 (1-419)

Alignment Scores:  
 Pred. No.: 1,48e-127 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-09-763-153-1 (1-419)

```

1  GCCAACTCCTCTCTGAGGAGCTCCGTACACAGCCTGGAGCGGAGTGCATAGAGGAG 60
1  AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 20
61  ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120
21  IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40
121  TGTCTCAAGCACATCGACGGTGACAGTGTGGTCTGGCTTCCCTTGGAGCACCGCTGGCC 180
41  TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60
181  AGCTGTGCTCGCGGACGCGACGTCATCGACGCGATCGGAGCTTCAGCTCGCACTGC 240
  
```

```
Db 61 SerLeuCysCysGlyHisGlyThrCysLeuAspGlyLeuGlySerPheSerCysAspCys 80
QY 241 CCACGGCTGGAGGGCGGCTTCTGCGAGCGGAGTGAGTCTCTCAATTCCTGCTGCTG 300
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100
QY 301 GACACGGCGCTGCACGCATCTACTGCTAGAGGAGTGGCTGGCGGCGCTGACTGT 360
Db 101 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgCysSerCys 120
QY 361 GCGCTGCTACAAGCTGGGGAGCGACTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420
Db 121 AlaProGlyTrpLysLeuGlyAspLeuGluCysHisProAlaValLysPhePro 140
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAACGCGGAGTCACTGAAACGAGACAGAA 480
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160
QY 481 GACCAAGAACCAAGTAGATCGCGCTCATTTGATGGGAAGATGACACGCGGGGAGAC 540
Db 161 AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgArgGlyAsp 180
QY 541 AGCCCTCGCAGTGTCTGCTGCTGACTCAAGAGAAAGCTGCTGCGGGGAGTGTCTC 600
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200
QY 601 ATCCACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 220
QY 661 GTGAGCTGGAGAGTATGACCTGGCGGCTGGGAGAGTGGGAGTGGAGTGGAGTGG 720
Db 221 ValArgLeuGlyLysTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 240
QY 721 AAGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAla 260
QY 781 CTGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysLeuProAsp 280
QY 841 AGCGGCTTGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGly 920
QY 901 GGCTACACAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 301 GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsn 320
QY 961 ATCAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetVal 340
QY 1021 GAGAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 341 GluAsnMetLeuCysAlaGlyLysLeuGlyAspArgGlnAspAlaCysGluGlyAsp 360
QY 1081 GGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSer 380
QY 1141 GGTGAGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 400
QY 1201 GACTGGATCCATGGGCGACATCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAG 1245
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415
```

RESULT 3

US-10-182-263-2

```
; Sequence 2, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-2
```

## Alignment Scores:

```
Pred. No.: 1.49e-127 Length: 461
Score: 2298.00 Matches: 415
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.33% Indels: 0
DB: 9 Gaps: 0
```

US-09-763-153-1 (1-1245) x US-10-182-263-2 (1-461)

```
QY 1 GCGAATCTCTCTGAGGAGCTCGCTCACAGCAGCTGGAGCGGAGTGCATGAGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 62
QY 61 ATCTGTGACTTCGAGAGAGCCCAAGAAATTTCCAAAATGTGGATGACACACTGCCTTC 120
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCAAGCAGCTCGACGGTGCAGCAGTGTGCTTCTGCTTGGAGCAGCCGCTGCGCC 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCTGTGCTGGCGCAGCGGACGCTGATCGACGCGATCGGCGAGTTCAGCTGCGACTGC 240
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGCTGGAGGGCGGCTTCTGCCAGCGAGTGCAGTGCAGTGCCTCAATTGCTCGCTG 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACACGGCGCTGCACGCATTTACTGCTAGAGGAGTGGGCTGGCGGCGCTGACTGT 360
Db 143 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgCysSerCys 162
QY 361 GCGCTGCTACAAGCTGGGGAGCGACTCTCTGAGTGTCAACCCGCGCTGAAAGTTCCTC 420
Db 163 AlaProGlyTrpLysLeuGlyAspLeuGluCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAACGCGGAGTGCAGTGCAGTGCAGTGCAGTGC 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAACCAAGTAGATCGCGGCTCATTTGATGGGAAGATGACACGCGGGGAGAC 540
Db 203 AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgGlyAsp 222
QY 541 AGCCCTCGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242
QY 601 ATCCACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 660
```

DB: 9 Gaps: 0  
US-09-763-153-1 (1-1245) x US-09-978-917A-2 (1-461)  
QY 1 GCCAACTTCTCTGAGGAGTCCCTACAGCAGCTGGAGCGGAGTGCATAGAGAG 60  
DB 43 ALaAsnSerPheLeuGluGluArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTACTTCCAGGAGGCAAGGAAATTTTCCAAATGTGGATGACACACTGCCCTTC 120  
DB 63 IleCysAspPheGluGluAlaLysGluIlePheGluAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAAGCAGCTCGACGCTGACAGTGGTCTTGGCTTGGCCCTGGAGCACCCTGCC 180  
DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 191 AGCTGTGCTGCGGCGCAGCGCAGCTGCATCGACGGCATCGGAGCTTCAGCTGCCAGTC 240  
DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGTTCCTCAATTCCTGCTG 300  
DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACACGCGCGCTGCACGCTTACTGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTGT 360  
DB 143 AspAsnGlyGlyCysThrHisGlyThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTGCTACAGCTGGGCGACGCTCTGTCAGTGTGCACCCCGAGTGAAGTTCCT 420  
DB 163 AlaProGlyTrpLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAAGCGGATGGAGAGAGCGCAGTACCTGAAACGAGACACAGAA 480  
DB 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGTCCGGGCTCATTTGATGGAAAGATGACACGCGGGAGAC 540  
DB 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyValaValLeu 242  
QY 601 ATCCACCTCTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTCAGGCTTGGAGAGTATGACCTGCGGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
DB 263 ValArgLeuGlyGlyLysArgLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGAGGTCTTCTCCACCCCACTACAGCAAGAGCACCACCGACCAATGACATCGCAGTC 780  
DB 283 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACCTGCGCCAGCGCGCCACCTCTCTGACAGCATAGTGCCTGCTGCTGCTGCTGCT 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTGCAGAGCGCGAGCTCAATCAGGCGCGCGCAGAGAGCTCTGACGGGCTGG 900  
DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACCAAGCAGCGCGAGAGAGGAGGCGCAAGAAACCGCAGCTTCCTCAACTTC 960  
DB 343 GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 363 IleLysIleProValProAlaProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTCAGGCTTGGAGAGTATGACCTGCGGCGCTGGGAGAAAGTGGAGCTGGACCTGGACATC 720  
DB 263 ValArgLeuGlyGlyLysArgLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGAGGTCTTCTCCACCCCACTACAGCAAGAGCACCACCGACCAATGACATCGCAGTC 780  
DB 283 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACCTGCGCCAGCGCGCAGCTCTCGCAGACCAATGAGTGCCTGCTGCTGCTGCTGCT 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTGCAGAGCGCGGAGCTCAATCAGGCGCGCGCAGAGAGCTTCCTGAGCGGCTGG 900  
DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACCAAGCAGCGCGAGAGAGGCGCAAGAAACCGCAGCTTCCTCAACTTC 960  
DB 343 GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 363 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

US-09-978-917A-2  
; Sequence 2, Application US/09978917A  
; Publication No. US20030027299A1  
; GENERAL INFORMATION:  
; APPLICANT: Maxygen Aps; Maxygen Holdings  
; TITLE OF INVENTION: Protein C or activated protein C-like molecules  
; FILE REFERENCE: 0219us310 - protein C  
; CURRENT APPLICATION NUMBER: US/09/978,917A  
; CURRENT FILING DATE: 2001-10-17  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 461  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(42)  
; FEATURE:  
; NAME/KEY: CHAIN  
; LOCATION: (43)...(461)  
US-09-978-917A-2  
Alignment Scores:  
Pred. No.: 1.49e-127 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0

Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGGTCCCTCCCTCCACGGCAGCTGGTTCCTGGTGGGCTGGTGGAGCTGG 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGCTCTTCCACAACTACGGCGTTTACACCAAGTCAGCCGCTAGCTC 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrglyValTyThrLysValSerArgTyLeu 442  
QY 1201 GACTGATCCATGGGCATCAGACAAAGAAAGCCGCCAGAAAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 5  
US-10-182-263-5  
; Sequence 5, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; PRIORITY FILING DATE: 2002-07-22  
; PRIOR FILING DATE: 2002-02-11  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-182-263-5

Alignment Scores:  
Pred. No.: 6,48e-126 Length: 419  
Score: 2270.00 Matches: 410  
Percent Similarity: 99.28% Conservative: 2  
Best Local Similarity: 98.80% Mismatches: 3  
Query Match: 97.13% Indels: 0  
DB: 9 Gaps: 0  
US-09-763-153-1 (1-1245) x US-10-182-263-5 (1-419)

QY 1 GCCAACTCCCTTCCTGGAGAGTCCGTCACAGCCCTGGAGCGGGAGTGATAGAGGAG 60  
Db 1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGluGlu 20  
QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTTCCAAATGTGGATGACACACTGGCCCTTC 120  
Db 21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspThrLeuAlaPhe 40  
QY 121 TGTCTCAAGCAGCTGCAGGCTGACAGTCTGGTCTTGGCTTGGCCCTGGAGCACCCGCTGGCC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCCTGTGTGGGGCAGCGCACGTCATCGACGGCATCGGCGAGCTTCAGCTGCAGCTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGGGCTGGAGGGCGGCTTCGCCAGCGGAGGTGAGTTCCTCAATGCTCGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAGCGGCTGCAGCATTTACTGCTAGAGAGGTGGCTGGCGGCTGCTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgCysSerCys 120

QY 361 GGCCTGGCTACAAGCTGGGGACGACCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420  
Db 121 AlaProGlyTyrlsLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGGCCCTCGAAGCGGATGGAGAGAGCGACGTACCTGAAACAGAGACACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgSerThrGlu 160  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACAGCGGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTGGCAGGTGGTCTCTGCTGACTCAAGAGAGAGTGGCTCGGGGGAGTGGCTC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTCTGGTGTCTGACAGCGGCCCACTGCATGGATGATGATCAAGAGCTCTT 660  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220  
QY 661 CTCAGGCTTGAGAGATGATGCTCGCGGCTGGGAGAGTGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGluTyArgPheArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGAGGTCTTCTCCACCCCACTACAGCAAGAGACACCCAGCATGATGATGATGATGATG 780  
Db 241 LysGluValPheValHisProAsnTyrlsSerLysSerThrThrAspAsnAspIleAlaLeu 260  
QY 781 CTGCACTGGCCCGCAGCGCCCTCTCCAGACCATAGTGCCTGCTGCTGCTGCTGCTGCTG 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
QY 841 AGCGGCTTCGAGAGCGGAGCTCAATAGCGCGGCGGAGAGACCTCTGAGCGGGCTGG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACCACAGCAGCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 301 GlyTyrlsSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGGTCCCGCACAATGATGTCAGGAGGTCATGATGATGATGATGATGATG 1020  
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GACAACATGCTGTGCTGGGAGTCTCTGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCCATGGTTCCTCTCCAGCGCACCTGGTTCCTGGTGGGCTGGTGGAGCTGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTTCCAACTACGGCGTTTACACCAAGTTCAGCGGCTGCTG 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrglyValTyThrLysValSerArgTyLeu 400  
QY 1201 GACTGATCCATGGGCATCAGACAAAGAAAGCCGCCAGAAAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

RESULT 6  
US-10-182-263-3  
; Sequence 3, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/181948



Tue Mar 18 16:19:31 2003

```

, PRIORITY DATE: 2002-02-11
, PRIORITY APPLICATION NUMBER: 60/189199
, PRIORITY FILING DATE: 2000-03-14
, NUMBER OF SEQ ID NOS: 12
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 3
, LENGTH: 419
, TYPE: prt
, ORGANISM: Homo sapiens
US-10-182-263-3

```

Alignment Scores:	1.46E-125	Length:	419
Pred. No.:	Score:	Matches:	409
Percent Similarity:	99.04%	Conservative:	2
Best Local Similarity:	98.55%	Mismatches:	4
Query Match:	96.88%	Indels:	0
DR:	9	Gaps:	0

US-09-763-153-1 (1-1245) X US-10-182-263-3 (1-419)

1	GCCAACTCCTCTCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGAG	60
1	AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGluGlu	20
61	ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATATGTGGATGACACACTGGCCCTC	120
21	IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspAspThrLeuAlaPhe	40
121	TGGTCCAAAGCACGTCGACGAGTACCAGTGCCTTCCCTTGGAGCACCGCTGCCCC	180
41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	60
181	AGCTGTGCTCGCGGCACGGCAGCTGCATCGACGGCATCGCAGCTTCACGTCGCACATGC	240
61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
241	CGCAGCGCTGGGAGGCGCGTCTGCCACGCGGAGGTGAGCTTCCTCAATTGCTCGCTG	300
81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
301	GACACGCGCTGCACGCATTACTGCTAGAGAGGTGGCTGGCGGCGCTGTAGCTGT	360
101	AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgCysSerCys	120
361	CGGCTGTGCTACAAGCTGGGGACGACCTCCTCGAGTGTCAACCCCGAGTGAAATTCCT	420
121	AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro	140
421	TGTGGAGGCCCTGGAACGGATGGAGNAGCGCAGTCACCTGAAACGAGACACAGAA	480
141	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	160
481	GACCAAGAACCAAGTAGATCCGCGGCTCATTCATGGGAGAGTACACAGCGCGGAGAC	540
161	AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp	180
541	AGCCCTGTGCAGGTGGTCCTGCTGGACTCAAGAAAGAGCTGGCTCGGGGGCAGTGCCTC	600
181	SerProTrpGlnValValLeuLeuAspSerLysLysLysSerAlaCysGlyAlaValLeu	200
601	ATCCACCCCTCTCTGGTGCTGACAGCGGCCACATGCATGATGATGATCCCAAGAAGCTCCTT	660
201	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	220
661	GTCAGCTTGGAGAGTATGACCTCGCGCGCTGGGAGAGTGGGAGCTGCACCTGGACATC	720
221	ValArgLeuGlyGlyTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle	240
721	AAGGAGGTCTTGTCACACCCCAACTACAGCAAGAGCACCACCCAGCATGATCATCGCACTG	780
241	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAlaIleAlaLeu	260

## RESULT 7

```

US-10-182-263-6      , Application US/10182263
, Sequence 6,
, Publication No. US200300223541
, GENERAL INFORMATION:
, APPLICANT: Gerlitz, Bruce E
, APPLICANT: Jones, Bryan E
, APPLICANT: Grinnell, Brian W
, TITLE OF INVENTION: PROTEIN C DERIV
, FILE REFERENCE: X-13611
, CURRENT APPLICATION NUMBER: US/10/
, CURRENT FILING DATE: 2002-07-22
, PRIOR APPLICATION NUMBER: 60/18194
, PRIOR FILING DATE: 2002-02-11
, PRIOR APPLICATION NUMBER: 60/18919
, PRIOR FILING DATE: 2000-03-14
, NUMBER OF SEQ ID NOS: 12
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 6
, LENGTH: 419
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-182-263-6

```

Alignment Scores:	1.91e-125	Length:	419
Pred. No.:	2262.00	Matches:	409
Score:		Conservative:	2
Percent Similarity:	99.04%	Mismatches:	0
Best Local Similarity:	98.55%	Indels:	0
Query Match:	96.79%	Gaps:	0
	9		

US-00-763-153-1 (1-1245) x US-10-182-263-6 (1-419)

QY	1	GCCAACTCCTTCTCGAGGAGCTCCGCTCACACAGCCCTGGAGCGGGAGTGCATAGAGGAG	60
Db	1	AlaAsnSerPheLeuGluGluLeuArGlnGlySerLeuGluArGluGluCysIleGluGlu	20
QY	61	ATCTGTGACTTCGAGGAGGCCAAGAAATTTTCCAAATGTGGATGCACACACTGGCCTTC	120
Db	21	IleCysAspPheGluAlaLysGluIlePheGluAspValAspPThrLeuAlaIlePhe	40

QY 121 TGGTCCAAAGCAGCTCGACGGTACCAGTGGCTTGGTCTTGGCCCTGGAGCACCCTGGCC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCCTGTGCTGGGGCAGCGACGTGCATCGACGCGATCGCGAGCTTCAGCTGGACTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCCAATTCGCTGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAAGCGGTGCACGCAATTAATCTAGAGAGGTGGCTGGCGGCGCTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCTGTGCTACAGCTGGGGAGCGACCTTCCTGCGAGTGTACCCCGCAGTGAAGTTCCT 420  
Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAGAGCGACCTACCTGCAACAGACACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAACCAAGTAGATCCCGCTCATTTGATGGAGAGATGACGAGCGGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTGGAGGTGGTCTGCTGGACTCAAGAGAGTGGCGTGGCGGCGAGTGC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCCTGGGTGCTGACAGCGCGCCACTGCTGATGATGAGTCCCAAGAGCTCCT 560  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 220  
QY 661 GTCAGGCTTGAGAGTAGTACCTCGCGCTGGGAGAGTGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGlyTrpAspLeuArgArgTrpGluLysTrpLeuAspLeuAspIle 240  
QY 721 AAGAGGTCTTGGTCCACCCCACTACAGCAAGAGCACCAGCAATGACATCGCAGT 780  
Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrIleValProLysCysLeuProAsp 260  
QY 781 CTGACCTGGCCCGCCACCCCTCTCGCAGACCATAGTGCACATCTGCTCCCGGAC 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysCysLeuProAsp 280  
QY 841 AGCGGCTTCGAGCGCGAGCTCAATCAGCGCGCGCCAGGAGACCTCTGACGGGCTGG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGAGAGAGGAGGCGCAAGAGAAACCGCACCTTCTCTCAACTC 960  
Db 301 GlyThrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGCTGCCGCAATAGTGCAGCGAGTGCATGAGCAACATGTGTCT 1020  
Db 321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGCTGTGCGGGCATCTCGGGGACCGGCGAGAGTGCCTGGAGGGCGACAGT 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCCATGCTGCTCTTCCAGGCGACCTGCTTCCGCTGGGCTGGTGGCTGGTGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTCACAACTACGGCGTTTACACAAAGTCAGCGCTACCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 400

QY 1201 GACTGTATCCATGGCACATCAGACAAAGGAAGCCCCCAGAAAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 8

US-10-182-263-4  
; Sequence 4, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Grimell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; PRIOR FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/181948  
; PRIOR FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/189199  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-182-263-4

## Alignment Scores:

Pred. No.:	2,5e-125	Length:	419
Score:	2260.00	Matches:	408
Percent Similarity:	99.04%	Conservative:	3
Best Local Similarity:	98.31%	Mismatches:	4
Query Match:	96.71%	Indels:	0
DB:	9	Gaps:	0

US-09-763-153-1 (1-1245) x US-10-182-263-4 (1-419)

QY 1 GCCAACTCTCTTCTGGAGGAGCTCCGTCACAGCAGCGCTGGAGGGGAGTGCATAGAGAG 60  
Db 1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGluGlu 20  
QY 61 ATCTGTCACTTCGAGGAGGCGCAAGGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
Db 21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspAspThrLeuAlaPhe 40  
QY 121 TGTCTCAAGCAGCTCGACGCTGACAGTGTCTGCTTGGCTGGAGCACCCTGGCGC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCCTGTGCTGGCGGACCGCAGCTGCATCGAGCGCATCGGAGCTTCAGCTGCGACTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCCAATTCGCTGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAAGCGGTGCACGCAATTAATCTAGAGAGGTGGCTGGCGGCGCTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCTGTGCTACAGCTGGGGAGCGACCTTCCTGCGAGTGTACCCCGCAGTGAAGTTCCT 420  
Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAGAGCGACCTACCTGCAACAGACACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAACCAAGTAGATCCCGCTCATTTGATGGAGAGATGACGAGCGGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180

Score:	736.00	Matches:	150
Percent Similarity:	52.36%	Conservative:	72
Best Local Similarity:	35.38%	Mismatches:	156
Query Match:	31.49%	Indels:	46
DB:	9	Gaps:	10

US-09-763-153-1 (1-1245) x US-10-132-829-5 (1-461)

QY	13	CTGAGAGAGCTCCGTCACACAGCAGCTGGAGCGGAGTGTAGAGGAGATCTGTGACTTC	72
DB	52	LeuGluGluPheValGlnGlyAsnLeuGluArgGluCysMetGluGluLysCysSerPhe	71
QY	73	GAGGAGCGCAAGAAATTTCCAAATGTGGATCACACACTGGCCCTTGTGGTCCAGACAC	132
DB	72	GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGlnTyr	91
QY	133	CTGACGGGTGACAGTGTCTGGTCTTCCCTTGGAGCACCCTGGCCGAGCTGTGCTGC	192
DB	92	ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu	103
QY	193	GGGACGCGCATCGATCGAGCGCATCGCAGCTTCAGTGTGCGACTGCGCGAGCGGTGG	252
DB	104	AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe	123
QY	253	GAGGCGCGCTTCTGCCAGCGCGAGGTGAGCTTCTCAATTTGCTGCTGGACACAGCGCGC	312
DB	124	GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg	140
QY	313	TGCAGCGATTACTGCCTAGAGGAGGTGGCTGGCGG---CGCTGTAGTGTGCGCTGGC	369
DB	141	CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly	160
QY	370	TACAAGCTGGGGGACGACCTCTGCTGCTCACCCCGCAGTGAAGTTCCCTTGTGGGAGG	429
DB	161	TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg	180
QY	430	CCCTGGAGCGGATGGAGAAAGCGGAGTCCCTGAAACGA-----	471
DB	181	ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProAspValAsp	200
QY	472	-----GACACAGAAAGACCAAGAACCAAGTAGAT-----	501
DB	201	TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer	220
QY	502	-----CCCGGCTCATGATGGGAGATGACCCAGCGGGGAGACAGCCCTGG	549
DB	221	PheAsnAspPheThrArgValValGlyGlyGluAspAlaLysProGlyGlnPheProTrp	240
QY	550	CAGGTGGTCTCTGCTGACTCAAGAAAGAGTGGCTGCGGGGAGTGCATCCACCCC	609
DB	241	GlnValVal-----LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu	259
QY	610	TCCTGGGTCTGACAGCGGCCCATGTCATGATGATGATGATGATGATGATGATGATGATG	669
DB	260	LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla	279
QY	670	GGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	729
DB	280	GlyGluHisAsnIleGluGluThrGluHisThrGluGlnLysArgAsnValIleArgIle	299
QY	730	TTGCTGCCCAACCACTACAGCAAGACACACAC-----GACAATGACATCGCTGCTG	783
DB	300	IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeu	319
QY	784	CACCTGGCGCCCGCCACCTCTCGCAGACCATAGTGGCCCATCTGCTCCCGACAGC	843
DB	320	GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys	339
QY	844	GGCTTGCAGAGCGCGAGCTCAATCAGCGCGCGAGGACACCTCTGACGCGGTGGGCG	903
DB	340	GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly	357
QY	904	-----TACCACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	957

QY	541	AGCCCTGGCAGGTGGTCTCTGCTGGACTCAAGAAAGAGTGGCTGCGGGGAGTGTCTC	600
DB	161	SerProTrpGlnValValLeuLeuAspSerLysLysSerAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTCTGGTGTGACAGCGGCGCCACTGATGATGATGATGATGATGATGATG	660
DB	201	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	220
QY	661	GTGAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
DB	221	ValArgLeuGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	240
QY	721	AAGGAGGCTTCTGCTCCACCCCACTACAGCAAGACACACCATGATGATGATGATGATG	780
DB	241	LysGluValPheValHisProAsnTyrSerLysSerThrSerAspAsnAspIleAlaLeu	260
QY	781	CTGCACTGGC	840
DB	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	280
QY	841	AGCGCGCTTGCAGAGCGCGAGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGC	900
DB	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACACAGAGCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	960
DB	301	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAGAGTCCGCTGCTCCGCACAACTGAGTGCAGCGAGGTCATGAGCAACATGCTGTCT	1020
DB	321	IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGACATGCTGTGCGGGCATCTCCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG	1080
DB	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GGGGGCGCATGCTGCTCTCCACGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
DB	361	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	380
QY	1141	GGTGAGGCTGTGGCTCTTCCACAACTACGCGGCTTTACACAAAGTACGCGGCTACCTC	1200
DB	381	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	400
QY	1201	GACTGATCATGCGGCACATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1245
DB	401	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	415

RESULT 9  
US-10-132-829-5  
Sequence 5, Application US/10132829  
Publication No. US20030044982A1  
GENERAL INFORMATION:  
APPLICANT: Chien, Kenneth R  
APPLICANT: Hoshijima, Masahiko  
TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII  
TITLE OF INVENTION: with vesicle vector  
FILE REFERENCE: 6627-PALL70  
CURRENT APPLICATION NUMBER: US/10/132,829  
CURRENT FILING DATE: 2002-04-25  
PRIOR APPLICATION NUMBER: 60/286,314  
PRIOR FILING DATE: 2001-04-25  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 461  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-829-5  
Alignment Scores: 5.21e-36 Length: 461  
Pred. No.:

Db 358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370  
 QY 958 TTCATCAAGATCCCGTGGTCCCGCACACATAGTGCAGCGAGTCAACACATGGTG 1017  
 Db 371 TyrLeuArgValProLeuValAspAlaThrCysLeuArgSerThrLysPheThrIle 390  
 QY 1018 TCTGAGAACATCTGTGTGCGGGCATCTCGGGACCGGAGGATGCTCGGAGGCGAC 1077  
 Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisIleGlyArgAspSerCysGlnGlyAsp 410  
 QY 1078 AGTGGGGGCCCATGCTGCTCCCTCCACGGCACCTGCTGTGCTGGGCTGGAGC 1137  
 Db 411 SerGlyGlyProHisValThrGluValGlyThrSerPheLeuThrGlyIleSer 430  
 1138 TGGGTGAGGCTGTGGGCTCTTCAACAACCTAGCGGCTTACACCAAGCTCAGCCCTAC 1197  
 431 TrpGlyGluGluCysAlaMetLysGlyTyrGlyIleTyrThrLysValSerArgTyr 450  
 QY 1198 CTCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454

## RESULT 10

US-09-884-901-3

; Sequence 3, Application US/09884901

; Patent No. US20020076798A1

; GENERAL INFORMATION:

; APPLICANT: Miao, Carol

; APPLICANT: Kay, Mark

; TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods of Use

; FILE REFERENCE: UOFW-1-17396

; CURRENT APPLICATION NUMBER: US/09/884,901

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: US 60/212,902

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 461

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-884-901-3

## Alignment Scores:

Pred. No.:	5,21e-36	Length:	461
Score:	736.00	Matches:	150
Percent Similarity:	52.36%	Conservative:	72
Local Similarity:	35.38%	Mismatches:	156
Query Match:	31.49%	Indels:	46
DB:	10	Gaps:	10

US-09-763-153-1 (1-1245) x US-09-884-901-3 (1-461)

QY 13 CTGAGAGCTCGTCCACAGCAGCTGGAGCGGAGTGCATAGAGGAGATCTGTGACTC 72  
 Db 52 LeuGluGluPheValGlnGlnLeuGluArgGluCysMetGluGluLysCysSerPhe 71  
 QY 73 GAGGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGCGCTGCTGCTGCTCCAGCAC 132  
 Db 72 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGlnTyr 91  
 QY 133 GTCAGCGGTGACAGTGTGCTGTGCTGCTTGGACACCGGTGCGGCGCTGTGCTGC 192  
 Db 92 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 103  
 QY 193 GGCACGCGCAGTCATCCACGGCATCGGCATTCAGCTGCGACTGCGCGAGCGGTGG 252  
 Db 104 AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 123  
 QY 253 GAGGCGCGCTTCTGCGCGCGGCGAGGTGAGCTTCTCTCAATTCGTGCTGGACACGCGGC 312  
 Db 124 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 140

## RESULT 11

US-09-118-748-2

; Sequence 2, Application US/09118748A

; Patent No. US20020031799A1

QY 313 TGCAGCGATTACTGCTAGAGAGGTGGTGGCGG---CGCTGACTGCTGGCGCTGGC 369  
 Db 141 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 160  
 QY 370 TACAAGCTGGGGGAGGACCTCTCAGTGTACCCCGCAGTGAAGTTCCTTGTGGGAGG 429  
 Db 161 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 180  
 QY 430 CCTGGAAGCGGTGGAGAAAGCGCAGTCCACCTGAACGA----- 471  
 Db 181 ValSerValSerGlnThrSerLysLeuThrArgAlaGluAlaValPheProAspValAsp 200  
 QY 472 -----GACACAGAACACCAAGAACCAAGTAGAT----- 501  
 Db 201 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer 220  
 QY 502 -----CGCGCGCTCATGTGAGGAAGATGACAGCGGGGAGGAGACCCCTGG 549  
 Db 221 PheAsnAspPheThrArgValValGlyGluAspAlaLysProGlyGlnPheProTrp 240  
 QY 550 CAGTGTGCTGCTGGACTCAAGAGAAAGCTGGCGGCGCAGTGTCTATCCACCCC 609  
 Db 241 GlnValVal---LeuAsnGlyLysValAspAlaPheCysGlySerIleValAsnGlu 259  
 QY 610 TCCTGGGTGTGACAGCGGCCACTGCTGATGATGATGCCAAGAGCTCCTGTGAGGCTT 669  
 Db 260 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla 279  
 QY 670 GGAGAGTATGACCTGGCGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGAGGTC 729  
 Db 280 GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 299  
 QY 730 TTGCTGCCACCCCACTACAGAACAGCACCCACCC-----GACAATGACATCCACTGTG 783  
 Db 300 IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeu 319  
 QY 784 CACCTGCCCGCCCGCCCGCCCTCTCGCAGACCATAGTCCCATGCTGCCGACAGC 843  
 Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAsp 339  
 QY 844 GGCCTTGCAGAGCGCGAGTCAATCAGGCGCGCCAGGAGACCCCTGTGACGGGTGGGC 903  
 Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 357  
 QY 904 -----TACCACAGCCCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957  
 Db 358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370  
 QY 958 TTCATCAAGATTCCTGCTGCTCCCGCACATGATGTCAGGAGGTGATGAGCAACATG 1017  
 Db 371 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 390  
 QY 1018 TCTGAGAACATCTGTGCTGGCGCATCTCCGGGACCGCAGGATGCTGCGAGGCGAC 1077  
 Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410  
 QY 1078 AGTGGGGGCCCATGCTGCTCCCTCTCCACGGCACCTGCTGCTGGGCTGGTGGAGC 1137  
 Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleSer 430  
 QY 1138 TGGGTGAGGCTGTGGGCTCTTCAACAACCTAGCGGCTTACACCAAGCTCAGCCCTAC 1197  
 Db 431 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 450  
 QY 1198 CTCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454



Best Local Similarity: 36.64% Mismatches: 157  
 Query Match: 31.07% Indels: 36  
 DB: 9 Gaps: 10

US-09-763-153-1 (1-1245) x US-10-109-498-1 (1-406)

```

QY 1 GCAACTCTCTCTCGAGGAGCTCGCTCACAGCAGCTGGAGCGGAGTGCAATAGAGGAG 60
Db 1 AlaAsnAlaPheLeu*****LeuArgProGlySerLeu***Arg***CysLys***** 20
QY 61 ATCTGTGACTTCGAGGAGCCCAAGAAATTTCCAAAATGTGGATGACACACATGCGCCTC 120
Db 21 GlnCysSerPhe*****AlaArg***IlePheLysAspAla***ArgThrLysLeuPhe 40
QY 121 TGTCTCAACGAGCTGCAGCGTGACAGCTGCTGGCTTGGCTTGGACACCCGCGGCC 180
Db 41 TrpIleSerTyrSerAspGlyAspGlnCys*****AlaSer 52
QY 181 AGCCTGTGTGCGGCGACGGCAGCTGCATCGACGGCATCGCAGCTTCAGCTGCGACATGC 240
Db 53 SerProCysGlnAsnGlyGlySerCysLysAspGlnLeuGlnSerTyrIleCysPheCys 72
QY 241 CCGACGGCTGGAGGCGGCTTCTGCGCAG---CCGAGGTGAGCTTCTCAATTTGCTCG 297
Db 73 LeuProAlaPheGluGlyArgAsnCysGluThrHisLysAspAspGlnLeuIleCysVal 92
QY 298 CTGGACACGGCGGCTGACGACCTTACTGCTAGAGAGGTGGCTGGCGCGC---TGT 354
Db 93 AsnGlnAsnGlyGlyCysGluGlnTyrCysSerAspHisThrGlyThrLysArgSerCys 112
QY 355 AGCTGTGCGCTTGGCTACAACTGGGGGACGACCTCTGCTGCTGCTGCTGCTGCTGCTG 414
Db 113 ArgCysHisGluGlyTyrSerLeuAlaAspGlyValSerCysThrProThrValGlu 132
QY 415 TTCCCTTGTGGAGGCGCTTGAAGCGGTGAGAGAGCGCAGTGCATGCTGCTGCTGCTGCTG 474
Db 133 TyrProCysGlyLys---IleProIleLeuGluLysArgAsnAla----- 146
QY 475 ACAGAAGACCAAGAAGACCAAGTAGATCCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 534
Db 147 -----SerLysProGlnGlyArgIleValGlyLysValCysProLys 161
QY 535 GGAGACACCCCTCGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
Db 162 GlyGluCysProThrGlnValLeuLeuValAsnGlyAlaGlnLeu---CysGlyGly 180
QY 595 GTGCTCATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Db 181 ThrLeuIleAsnThrIleThrValValSerAlaAlaHisCysPheAspLysIleLysAsn 200
QY 652 -----AGCTCTCTGCTGAGAGTATGACCTGCGGCGCTGGGAGAGTGGGAG 705
Db 201 TrpArgAsnLeuIleAlaValLeuGlyGluHisAspLeuSerGluHisAspGlyAspGlu 220
QY 706 CTGGACCTGGACATCAAGGAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
Db 221 GlnSerArgValAlaGlnValIleProSerThrTyrValProGlyThrThrAsn 240
QY 766 AATGACATCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
Db 241 HisAspIleAlaLeuLeuArgLeuHisGlnProValValLeuThrAspHisValPro 260
QY 826 ATCTGCTCCCGACAGCGGCTTGGAGCGGAGCTCAATCAGCGCGGCGGAGAGAC 885
Db 261 LeuCysLeuProGluThrThrPheSerGluArgThrLeuAlaPheVal---ArgPheSer 279
QY 886 CTGCTGACGGGCTGGGCTACACAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 945
Db 280 LeuValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla----- 294
QY 946 TTGCTCTCAACTCATAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
Db 295 LeuGluLeuMetValLeuAsnValProArgLeuMetThrGlnAspCysLeuGlnSer 314

```

```

QY 1006 -----AGCACATGCTGTGTGAGAACATGCTGTGCGGCGCATCCTCGG 1050
Db 315 ArgLysValGlyAspSerProAsnIleThrGluTyrMetPheCysAlaGlyTyrSerAsp 334
QY 1051 GACCGGAGGATGCTCGAGGCGGACAGTGGGGGCGCCCATGTCGCTCTTCCACGCG 1110
Db 335 GlySerLysAspSerCysLysGlyAspSerGlyGlyProHisAlaThrHisTyrArgGly 354
QY 1111 ACCTGTTCTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1170
Db 355 ThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThrValGlyHisPhe 374
QY 1171 GCGCTTTACCAAGTCAGCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
Db 375 GlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeuMetArgSerGlu 394
QY 1231 GAAGCCGCC 1239
Db 395 ProArgPro 397

```

## RESULT 13

```

US-09-978-295A-169
; Sequence 169, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C11
; CURRENT FILING DATE: 2001-10-15
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641

```

[illegible]



; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

## Alignment Scores:

Pred. No.: 2.07e-20 Length: 802  
 Score: 470.00 Matches: 124  
 Percent Similarity: 46.93% Conservative: 52  
 Best Local Similarity: 33.07% Mismatches: 145  
 Query Match: 20.11% Indels: 54  
 Ds: 9 Gaps: 18

US-09-763-153-1 (1-1245) x US-09-978-295A-169 (1-802)

```

172 CGGTGCGCCAGC-----CTGTGCTGGCGGCGGCGGCGGCTTCTGCCAGCGCAG 216
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 ProCysProGlyGluPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 467
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
217 ATCGGCAGCTTACGCTGCGAGTCCGCGCGGCGGCGGCGGCGGCTTCTGCCAGCGCAG 276
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
468 ValLys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg 483
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277 GTGAGCTTCTCCTAATGCTCGTGAGCAAGCGGCGGCTGACGATTTACTGCTAGAGGAG 336
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
484 AlaThrPhe-----GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 GTGGCTTGGCGGCGCTAGCTGCGCTGCGCTGCTACAACTGGGGGAGCACTCTCTGCGAG 396
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGluGln 515
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
397 TGTACACCCCGAGTGAAGTTCCTTGTGGGAGGCGGCTGGAAGCGGATGAGAAAGAGCGC 456
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
457 AGTCACCTGAAA-----CGAGACACAGAA 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
533 SerCysValLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 GACCAAGAA-----GACCAAGTAGATCCCGGCTCATTTGATGGGAG 522
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
553 AspGluGluHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyAla 572
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
523 ATGACGAGCGGGGAGACAGCCCTGCGAGGTGGTCTGCTGGACTCAAGAAAGAAAGCTG 582
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
573 ValSerSerGluGlyGluTrpProTrpGln---AlaSerLeuGlnValArgGlyArgHis 591
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 GCGTGGGGGCGAGTGTCTATCCACCCCTCTGGGTCGTCAGAGCGGCGGCGGCTGCTGAT 642
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
643 GAGTCCAAAGAGCTCTGTGAGGCTTGGAGAGTATGACCTGCGGCGG---TGGGAG--- 696
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
697 ---AAGTGG-----GAGTGGACCTGGACATCAAGAGGTCTTCGTCACCCCACTAC 747
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTyHis 651
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 AGCAAGAGCACACCGACCAATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 GluGluAspSerHisAspTyAspValAlaLeuLeuGlnLeuAspHisProValValArg 671
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 TCAGACACCATAGTCCCATCTGCTCCCGGAGAGCGGCTTGGAGAGCGCGAGCTCAAT 867
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 CAGGCGCGCGGAGAGACCTCGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGly 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
928 GCCAAGAAACCGGCACTCTGCTCCTCACTTCAATCAAGATTCGCTGGTCCCGGCAAT 987
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722
QY 988 GAGTGCACGAGGTCATGAGCAACATGGTGTGTGAGAACATGTGTGTGCGGCGCATCCTC 1047
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 LeuCysSerGluAlaTyArgTyrGlnValThrProArgMetLeuCysAlaGlyTyrArg 742
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1048 GGGACCGCGGAGGATGCTGCGAGGCGAGAGTGGGGGCGGCGGCGGCGGCGGCGGCGG 1104
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 743 LysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1105 CACGGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 782
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1165 AACTACCGCGGTTTACACCAAGTACCGCGGCTACTCGACTGGGATC 1209
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 TyrPheGlyValThrArgIleThrGlyValIleSerTrpIle 797

```

## RESULT 14

US-09-978-697-169  
 ; Sequence 169, Application US/09978697  
 ; Patent No. US20020169284A1

## GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kijavlin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2630P1C27  
 ; CURRENT APPLICATION NUMBER: US/09/978,697  
 ; CURRENT FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11

us-09-763-153-1.n2p.rapb

PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704





pred. NO.:	2.07e-20	Length:	802
Score:	470.00	Matches:	124
Percent Similarity:	46.93%	Conservative:	52
Best Local Similarity:	33.07%	Mismatches:	145
Query Match:	20.11%	Indels:	54
	9	Gaps:	18
DB:			

	ProCysProGluGluPheLeuCysSerValAsnGlyLeuCysValProAlaCysAspGly	467
448	ProCysProGluGluPheLeuCysSerValAsnGlyLeuCysValProAlaCysAspGly	467
217	ATCGGCAGCTTCAGCTGGCACTGCCGACGCGGTGGGAGGCGCGTCTCCAGCGGAG	276
468	ValLys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg	483
277	GTGAGCTCTCTCAATTGCTGCTGGACAACGGCGTGACGCATTACTCGCTAGAGGAG	336
484	AlaThrPhe-----GlnCysIysGluAspSer-----ThrCysIleSerLeuProLys	499
337	GTGGGCTGGCGGCTGTAGCTGTGGCTGGCTGCTACAAGCTGGGGGAGCACCTCTCGAG	396
500	Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln	515
397	TGTCAACCCCAGTAGAAGTTCTTGTGGAGGCGCTGGAGCGCATGGAGAAGAAGCGC	456
516	CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg	532
457	AGTCACCTGAATA-----CGACACACAGAA	480
533	SerCysValIysLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer	552
481	GACCAAGAAA-----GACCAGTAGATCGCGGCTCATTCATGGGAG	522
553	AspGluGluHisCysAspCysGlyLeuGlnGlyProSerSerArgileValGlyGlyAla	572
523	ATGACAGCGGGGAGACAGCCCTCGGCAGGTGGTCTGCTGACTCAACAAGAAGAGCTG	582
573	ValSerSerGluGlyGluTrpProTripeIn---AlaSerLeuGlnValArgGlyArgHis	591
583	GCTGTGGGGGAGTCTCATCCACCCCTCGGTGGTGTGACAGCGGCCCATCATGGAT	642
592	IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln	611
643	GAGTCCAAGAAGCTCCTCTGTACGGCTTGGAGATGATGACTCGCGGGC---TGGGAG---	696
612	GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn	631
697	---AAGTGG-----GAGCTGACCTGGACATCAAGGAGGTCTTCGTCCACCCCACTAC	747
632	SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTyris	651
748	AGCAAGAGCACCCAGCAATGACATCGCACTGCTGCACCTGGCCCCAGCCGCCACCTC	807
652	GluGluAspSerHisAspTyAspValAlaLeuLeuGlnLeuAspHisProValValArg	671
808	TCGAGAGACCATAGTGCCTCGCTCCCGGACAGCGGCTTGTGACGGCGGAGTCAAT	867
672	SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe	687
868	CAGCGCGCGGAGGAGCCCTCGTCAGCGGTGGGGCTTACACAGACCGCAGAGAAGGAG	927
688	GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly	705
928	GCCAAAGAGAAACCGCACCTTCGTCTCACTTCATCAAGATTCCCGTGGTCCCGCACAAT	987
706	ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp	722
988	GAGTCAGCGAGGTGATGAGCAACATGTGTGCTGAGAACAFGTGTGTGGGGGCATCCTC	1047
723	LeuCysSerGluAlaTyArgTrpGlnValThrProArgMetLeuCysAlaGlyTyArg	742

1	PRIOR FILING DATE: 1998-04-22
2	PRIOR APPLICATION NUMBER: 60/082796
3	PRIOR FILING DATE: 1998-04-23
4	PRIOR APPLICATION NUMBER: 60/083336
5	PRIOR FILING DATE: 1998-04-27
6	PRIOR APPLICATION NUMBER: 60/083322
7	PRIOR FILING DATE: 1998-04-28
8	PRIOR APPLICATION NUMBER: 60/083392
9	PRIOR FILING DATE: 1998-04-29
10	PRIOR APPLICATION NUMBER: 60/083495
11	PRIOR FILING DATE: 1998-04-29
12	PRIOR APPLICATION NUMBER: 60/083496
13	PRIOR FILING DATE: 1998-04-29
14	PRIOR APPLICATION NUMBER: 60/083499
15	PRIOR FILING DATE: 1998-04-29
16	PRIOR APPLICATION NUMBER: 60/083545
17	PRIOR FILING DATE: 1998-04-29
18	PRIOR APPLICATION NUMBER: 60/083554
19	PRIOR FILING DATE: 1998-04-29
20	PRIOR APPLICATION NUMBER: 60/083558
21	PRIOR FILING DATE: 1998-04-29
22	PRIOR APPLICATION NUMBER: 60/083559
23	PRIOR FILING DATE: 1998-04-29
24	PRIOR APPLICATION NUMBER: 60/083500
25	PRIOR FILING DATE: 1998-04-29
26	PRIOR APPLICATION NUMBER: 60/083742
27	PRIOR FILING DATE: 1998-04-30
28	PRIOR APPLICATION NUMBER: 60/084366
29	PRIOR FILING DATE: 1998-05-05
30	PRIOR APPLICATION NUMBER: 60/084414
31	PRIOR FILING DATE: 1998-05-06
32	PRIOR APPLICATION NUMBER: 60/084441
33	PRIOR FILING DATE: 1998-05-06
34	PRIOR APPLICATION NUMBER: 60/084637
35	PRIOR FILING DATE: 1998-05-07
36	PRIOR APPLICATION NUMBER: 60/084639
37	PRIOR FILING DATE: 1998-05-07
38	PRIOR APPLICATION NUMBER: 60/084640
39	PRIOR FILING DATE: 1998-05-07
40	PRIOR APPLICATION NUMBER: 60/084598
41	PRIOR FILING DATE: 1998-05-07
42	PRIOR APPLICATION NUMBER: 60/084600
43	PRIOR FILING DATE: 1998-05-07
44	PRIOR APPLICATION NUMBER: 60/084627
45	PRIOR FILING DATE: 1998-05-07
46	PRIOR APPLICATION NUMBER: 60/084643
47	PRIOR FILING DATE: 1998-05-07
48	PRIOR APPLICATION NUMBER: 60/085339
49	PRIOR FILING DATE: 1998-05-13
50	PRIOR APPLICATION NUMBER: 60/085338
51	PRIOR FILING DATE: 1998-05-13
52	PRIOR APPLICATION NUMBER: 60/085323
53	PRIOR FILING DATE: 1998-05-13
54	PRIOR APPLICATION NUMBER: 60/085582
55	PRIOR FILING DATE: 1998-05-15
56	PRIOR APPLICATION NUMBER: 60/085700
57	PRIOR FILING DATE: 1998-05-15
58	PRIOR APPLICATION NUMBER: 60/085689
59	PRIOR FILING DATE: 1998-05-15
60	PRIOR APPLICATION NUMBER: 60/085579
61	PRIOR FILING DATE: 1998-05-15
62	PRIOR APPLICATION NUMBER: 60/085580
63	PRIOR FILING DATE: 1998-05-15
64	PRIOR APPLICATION NUMBER: 60/085573
65	PRIOR FILING DATE: 1998-05-15
66	PRIOR APPLICATION NUMBER: 60/085704
67	PRIOR FILING DATE: 1998-05-15
68	PRIOR APPLICATION NUMBER: 60/085697

Length:	802
Matches:	124
Conservative:	52





us-09-763-153-1.n2p.rapb

Tue Mar 18 16:19:31 2003

```

; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Alignment Scores:
Pred. No.: 2.07e-20
Score: 470.00
Percent Similarity: 46.93%
Best Local Similarity: 33.07%
Query Match: 20.11%
DB: 9
Length: 802
Matches: 124
Conservative: 52
Mismatch: 145
Indels: 54
Gaps: 18

```

US-09-763-153-1 (1-1245) x US-09-978-189-169 (1-802)

QY 172 CCGTGGCCGAGC-----CTGTGCTGGGCACGCGCATC-----GACGCC 216  
Db ||||| ||||| :||| :||| :||| :|||

QY 448 ProCysProGlyGluPheLeuCysValAsnGlyLeuCysValProAlaCysaspGly 467  
Db ||||| ||||| :||| :||| :||| :|||

QY 217 ATCGCGAGCTTCAGTCCGACTCCGCGAGCGGCTGGGCGGCTCTCTCCAGCGCGAG 276  
Db ||||| ||||| :||| :||| :||| :|||

QY 468 ValLys-----AspCysProAsnGlyLeuAspGlnArgSerHisPhePhe 483  
Db ||||| ||||| :||| :||| :||| :|||

QY 277 GTGAGCTTCCTCAATTTGCTGGACAACGCGCGCTGCACGCAATTACTCGTAGAGGAG 336  
Db ||||| ||||| :||| :||| :||| :|||

QY 484 AlaThrPhe---GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499  
Db ||||| ||||| :||| :||| :||| :|||

QY 337 GTGGGCTGGCGGCTGTAGCTGGCGCTGCTACAACTGGCGGAGAGCACTCTCTGAG 396  
Db ||||| ||||| :||| :||| :||| :|||

QY 500 Val-----CysaspGlyGlnProaspCysLeuAsnGlySerAspGluGln 515  
Db ||||| ||||| :||| :||| :||| :|||

QY 397 TGTACACCCGCGAGTTCCTGTGGAGCGCCCTGGAAGCGGATGGAGAAGAGCGC 456  
Db ||||| ||||| :||| :||| :||| :|||

QY 516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532  
Db ||||| ||||| :||| :||| :||| :|||

QY 457 AGTCACCTGAAA-----CGAGCACAGAA 480  
Db ||||| ||||| :||| :||| :||| :|||

QY 533 SerCysValLysProAsnProGlnCysaspGlyArgProaspCysArgaspGlySer 552  
Db ||||| ||||| :||| :||| :||| :|||

QY 481 GACCAGAA-----GACCAAGTAGATCCGCGCTCATGTGGGAAG 522  
Db ||||| ||||| :||| :||| :||| :|||

QY 553 AspGluGluHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyVala 572  
Db ||||| ||||| :||| :||| :||| :|||

QY 523 ATCAGCAGCGGGGAGACAGCCCTGGCGAGTGCTCTGGGACTCAAGAAGAACGCTG 582  
Db ||||| ||||| :||| :||| :||| :|||

QY 573 ValSerSerGluGlyGluTrpProArgPgn---AlaSerLeuGlnValArgGlyArgHis 591  
Db ||||| ||||| :||| :||| :||| :|||

QY 583 GCCTCGGGGAGTCTCATCCACCCCTCTCTGGTGTGACAGCGGCCCACTCATGGAT 642  
Db ||||| ||||| :||| :||| :||| :|||

QY 592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611  
Db ||||| ||||| :||| :||| :||| :|||

QY 643 GAGTCCAAGAAGCTCTTGTGAGCTGGAGATGATGACCTGGCGGC-----TGGAG--- 696  
Db ||||| ||||| :||| :||| :||| :|||

QY 612 GluaspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631  
Db ||||| ||||| :||| :||| :||| :|||

QY 697 ---AAGTG-----GAGCTGAGCTGGACATCAAGGAGTCTTCGTCACCCCACACTAC 747  
Db ||||| ||||| :||| :||| :||| :|||

QY 632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTyHis 651  
Db ||||| ||||| :||| :||| :||| :|||

QY 748 AGCAAGACACCCGACATGACATCGACTGCTGACCTGCCCGCCGCGCCACCTC 807  
Db ||||| ||||| :||| :||| :||| :|||

QY 652 GluLysSerHisAspTyAspValAlaLeuLeuGlnLeuAspHisProValValArg 671  
Db ||||| ||||| :||| :||| :||| :|||

QY 808 TCGCAGACCATAGTGCCTCTGCTCCCGACAGCGGCTTCGACAGCGCGAGCTCAAT 867  
Db ||||| ||||| :||| :||| :||| :|||

QY 672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687  
Db ||||| ||||| :||| :||| :||| :|||

QY 868 CAGCGCGCGGAGAGCCCTGTCGAGCGGCTGGGGTACACAGCAGCGGAGAGAGAG 927  
Db ||||| ||||| :||| :||| :||| :|||

QY 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly 705  
Db ||||| ||||| :||| :||| :||| :|||

QY 928 GCCAAGAGAAACCGCACCTTCGCTCCCTCACTTCATCAAGATTCCTGCTCCGCAAT 987  
Db ||||| ||||| :||| :||| :||| :|||

QY 706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722  
Db ||||| ||||| :||| :||| :||| :|||

QY 988 GAGTGCAGCGAGTCAATGACCAATGTTCTGAGAATCTGCTGTGGCGGATCCTC 1047  
Db ||||| ||||| :||| :||| :||| :|||

QY 723 LeuCysSerGluAlaTyArgTrpGlnValThrProArgMetLeuCysAlaGlyTyArg 742  
Db ||||| ||||| :||| :||| :||| :|||

QY 1048 GGGACCGCGGAGTCCCTGCGAGCGGACAGTGGGGGCCCATGTGTCG---TCCTTC 1104  
Db ||||| ||||| :||| :||| :||| :|||

QY 743 LysGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762  
Db ||||| ||||| :||| :||| :||| :|||

QY 1105 CAGCGACCTGTTCTTGTGGGCTGTGAGCTGGGTGAGGCTGTGGCTCTTCAC 1164  
Db ||||| ||||| :||| :||| :||| :|||

US-09-763-153-1 (1-1245) x US-09-978-189-169 (1-802)

QY 172 CCGTGGCCGAGC-----CTGTGCTGGGCACGCGCATC-----GACGCC 216  
Db ||||| ||||| :||| :||| :||| :|||

QY 448 ProCysProGlyGluPheLeuCysValAsnGlyLeuCysValProAlaCysaspGly 467  
Db ||||| ||||| :||| :||| :||| :|||

QY 217 ATCGCGAGCTTCAGTCCGACTCCGCGAGCGGCTGGGCGGCTCTCTCCAGCGCGAG 276  
Db ||||| ||||| :||| :||| :||| :|||

QY 468 ValLys-----AspCysProAsnGlyLeuAspGlnArgSerHisPhePhe 483  
Db ||||| ||||| :||| :||| :||| :|||

QY 277 GTGAGCTTCCTCAATTTGCTGGACAACGCGCGCTGCACGCAATTACTCGTAGAGGAG 336  
Db ||||| ||||| :||| :||| :||| :|||

QY 484 AlaThrPhe---GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499  
Db ||||| ||||| :||| :||| :||| :|||

RESULT 19  
US-09-888-615-113  
; Sequence 113, Application US/09888615  
; Patent NO. US20020064856A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: CAENPEPEL, SEAN  
; APPLICANT: CHARYDCZAK, GLEN

[illegible]

QY 808 TCAGAGACCATAGTCCCATCTGCTCCCGGACAGCGGCTTCAGAGCGCGAGCTCAAT 867  
 Db 672 SerAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687  
 QY 868 CAGCGCGCGGACAGAGACCTCTGAGCGGCTGGGCTACACAGAGCGCGAGAGAGAG 927  
 Db 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly 705  
 QY 928 GCCAAGAGAAACCGACCTTCCTCTCACTTCACTCAAGATTCCTCGGTGCCGACAAAT 987  
 Db 706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722  
 QY 988 GAGTCAGCAGGAGTCATGAGCAACATGCTGTGAGAACATCTGTGCGGGGCTCCTC 1047  
 Db 723 LeuCysSerGluAlaTyrArgTyrGlnValThrProArgMetLeuValCysAlaGlyTyrArg 742  
 1048 GGGGACCGCAGGATGCTGCGAGGGGACAGTGGGGGCGCCATGGTCCGC-----TCCCTTC 1104  
 Db 743 LysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762  
 QY 1105 CACGGCACCTGCTCTGCTGGGCTGCTGAGTGGGCTGGGCTGGGCTGCTCTTTCAC 1164  
 Db 763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 782  
 QY 1165 AACTACGGCGTTTACACCAAGTACCGCGCTACCTCGACTGGATC 1209  
 Db 783 TyrPheGlyValTyrThrArgIleThrGlyValIleSerTrpIle 797  
 RESULT 20  
 US-09-888-615-108  
 ; Sequence 108, Application US/09888615  
 ; Patent No. US20020064856A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PLOWMAN, GREGORY  
 ; APPLICANT: WHITE, DAVID  
 ; APPLICANT: CAENEPEEL, SEAN  
 ; APPLICANT: CHARYDCZAK, GLEN  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: NOVEL PROTEASES  
 ; FILE REFERENCE: 038602/1214  
 ; CURRENT APPLICATION NUMBER: US/09/888,615  
 ; PRIORITY FILING DATE: 2001-06-26  
 ; PRIOR APPLICATION NUMBER: 60/214,047  
 ; PRIOR FILING DATE: 2000-06-26  
 ; NUMBER OF SEQ ID NOS: 150  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 108  
 ; LENGTH: 850  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-888-615-108  
 Alignment Scores:  
 Pred. No.: 5,62e-18 Length: 850  
 Score: 428.50 Matches: 120  
 Percent Similarity: 43.99% Conservative: 52  
 Best Local Similarity: 30.69% Mismatches: 150  
 Query Match: 18.34% Indels: 69  
 DB: 10 Gaps: 16  
 US-09-763-153-1 (1-1245) x US-09-888-615-108 (1-850)  
 QY 163 TTGAGGACCGCTGTC-----GCCAGCCTGTGCTGGCGGACCGCAGCTGATC----- 210  
 Db 487 IleSerGlnProCysProValGlySerPheArgCysSerSerGlyLeuValProGln 505  
 QY 211 -----GACGGCATCGGCGAGCTTACGTGCGAGTCCCGCAGCGGCTGGGAGGC 258  
 Db 507 AlaGlnArgCysAspGlyValAsn-----AspCysPheAspGluSerAspGlu 522  
 QY 259 CGCTTCTGCCAGCGGAGGTGAGCTTCTCAATGTCTG----- 297

Db 523 LeuPheCysValSerProGlnProAlaCysAsnThrSerSerPheArgGlnHisGlyPro 542  
 QY 298 -----CTGACAAACCGCGGCTGCACGCATTAATCTGCTAGAG 333  
 Db 543 LeuIleCysAspGlyPheArgAspCysGluAsnGlyArgAspGluGlnAsnCysThrGln 562  
 QY 334 CAGTGGGCTGGCGGCTGTAGCTGTGGCTGGCTACAGCTGGGGGAGCGACCTC--- 390  
 Db 563 SerIleProCysAsnAsnArgThr-----PheLysCysGlyAsnAspIleCys 578  
 QY 391 -----CTGCACTGTCACCCCGCAGTGAAGTTCCTTCTGGAGGCCCTCG 435  
 Db 579 PheArgLysGlnAsnAlaLysCysAspGlyThrValAspCysProAspGly----- 595  
 QY 436 AAGCGGATGGAGAAGACGACGTCACCTGAACAGCAGACACAGAAAGAACACCAA 495  
 Db 596 -----SerAspGluGluGlyCysThrCysSerArgSerSerSer 608  
 QY 496 GTAGATCCGGCTCATTTGATGGGAAGATACACAGCGGGGAGACACCCCTGGCAGGTG 555  
 Db 609 AlaLeuHisArgIleIleGlyGlyThrAspThrLeuGluGlyGlyTrpProTrpGln--- 627  
 QY 556 GTCCTGCTGACTCAAAAGAAAGACTGCGCTGGCGGCGAGTGTCTATCCACCCCTCCTGG 615  
 Db 628 ValSerLeuHisPheValGlySerAlaTyrCysGlyAlaSerValIleSerArgGluTrp 647  
 QY 616 GTGCTGACAGCGGCCACTGC-----ATGGATGAGTCCAAAGAGCTCCTT 660  
 Db 648 LeuLeuSerAlaAlaHisCysPheHisGlyAsnArgLeuSerAspProThrProTrpThr 667  
 QY 661 GTCAGGCTTGAGAGATGATGACCTGCGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
 Db 668 AlaHisLeuGlyMetTyr---ValGlnGlyAsnAlaLysPheVal-----SerProVal 684  
 QY 721 AAGGAGCTCTGCTCCACCCCACTACAGCAAGAGACACCCAGCAATGATCATCGACTG 780  
 Db 685 ArgArgIleValValHisGluTyrTyrAsnSerGlnThrPheAspTyrAspIleAlaLeu 704  
 QY 781 CTGACACTG-----GCCAGCGCGCCACCTCTCGCAGACCATAGTGCCTCCTCCTC 834  
 Db 705 LeuGlnLeuSerIleAlaTrpProGluThrLeuLysGlnLeuIleGlnProIleCysIle 724  
 QY 835 CCGGACAGCGGCTTGCAGAGCGGAGCTCAATCAGCGCGCGGAGACCTCGTGAGC 894  
 Db 725 ProProThrGlyGlnArgValArg-----SerGlyGlyLysCysTrpValThr 740  
 QY 895 GGCTGGGCTACCCACAGCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954  
 Db 741 GlyTrpGly-----ArgArgHisGluAlaAspAsnLysGlySerLeuValLeu 756  
 QY 955 AACTTCATCAAGATCCCGTGGTCCCGCACAAATGAGTGCAGCGAGGTCTATGAGCAATG 1014  
 Db 757 GlnGlnAlaGluValGluLeuIleAspGlnThrLeuCysValSerThrTyr---GlyIle 775  
 QY 1015 GTCTCTCAGAACATGCTGTGCTGGCGCATCTCCGGGAGCGGAGAGATGCTCGCAGAGGC 1074  
 Db 776 IleThrSerArgMetLeuCysAlaGlyIleMetSerGlyLysArgAspAlaCysLysGly 795  
 QY 1075 GACAGTGGGGGGCCCATGGTGGCC-----TCCTTCCAGCGACCTGTTCTCGTGGGCG 1128  
 Db 796 AspSerGlyGlyProLeuSerCysArgArgLysSerAspGlyLysTrpIleLeuThrGly 815  
 QY 1129 CTGGTGGAGTGGGCTGAGGCTGTGGCTCTCTTACAACTAGCGGCTTTACACCAAGTTC 1188  
 Db 816 IleValSerTrpGlyHisGlyCysGlyArgProAsnPheProGlyValTyrThrArgVal 835  
 QY 1189 AGCCGCTACTCGACTGGATTCATGGGCACATC 1221  
 Db 836 SerAsnPheValProTrpIleHisLysTyrVal 846

Tue Mar 18 16:19:31 2003

us-09-763-153-1.n2p.rapb

Page 24

Job time : 66 secs